

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 12:51:18 ; Search time 446.127 Seconds  
(without alignments)  
11200.072 Million cell updates/sec

Title: US-09-745-763-35

Perfect score: 1851  
Sequence: 1 GGCTAGGCGCCGAGCTTAGT.....CTGAAAAAAAAAAAAAAAAAAAA 1851

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_19Jun03:.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1851	100.0	1851	19	AAV82779
2	1851	100.0	1851	24	ABO92016
3	1846.2	99.7	1863	21	AAZ98034
4	1846.2	99.7	1863	22	AAD11647
5	1846.2	99.7	1863	24	ABK65743
6	1846.2	99.7	1863	25	ACC50817
7	1846.2	99.7	1863	25	ABZ71453
8	1833.4	99.0	1923	21	AAZ98139

9	1830.8	98.9	1884	21	AAZ58313	Human peptidase NA
10	1830.2	98.9	2077	25	ACC50603	Human secreted pro
11	1830.2	98.9	2077	25	ABZ71331	Secreted protein-e
12	1820.4	98.3	1860	22	AAK94491	Human full-length
13	1751.2	94.6	1767	21	AAA40493	Human fetal kidney
14	1751	94.6	1778	18	AAV02296	Human secreted pro
15	1617.2	87.4	1895	22	AAH93703	Human protein enco
16	1491	80.6	1784	23	AAH73592	DNA encoding novel
17	1145	61.9	1778	24	ABK62598	Rat sequence diffe
18	1088.4	59.3	1134	21	AAZ28101	Human secreted pro
19	1098.4	59.3	1134	21	AAD11714	Human secreted pro
20	1098.4	59.3	1134	24	ABK69810	Human secreted pro
21	1098.4	59.3	1134	25	ACC50818	Human secreted pro
22	1098.4	59.3	1134	25	ABZ71454	Secreted protein-e
23	865.8	46.8	895	21	AAA44369	Human secreted exp
24	569	30.7	609	20	AAV87487	EST clone BU45. H
25	568.2	30.7	642	22	AAK92181	Human CDNA 5'-end
26	568.2	30.7	642	22	AAK93447	Human CDNA clone r
27	410.6	22.2	492	22	AAK92932	Human CDNA 3'-end
28	375.8	20.3	394	14	AAO61260	Human brain Expres
29	375	20.3	424	23	AA573591	DNA encoding novel
30	324.4	17.5	360	18	AAV88078	3' portion of CDNA
31	324.4	17.5	360	18	AAV02139	Human secreted pro
32	324.4	17.5	360	18	AAV02140	Human secreted pro
33	310.2	16.8	317	20	AAK40587	Human secreted pro
34	290.8	15.4	462	18	AAV86369	EST clone AM282.
35	284.6	15.4	462	18	AAV797398	Human secreted pro
36	284.6	15.4	462	18	AAV88060	Partial CDNA clone
37	284.6	15.4	462	18	AAV02139	Human secreted pro
38	277.8	15.0	441	22	AAK56831	Human immune/haema
39	233	12.6	234	24	ABN97386	Gene #3884 used to
40	122.4	6.6	217	20	AAK40583	Human secreted pro
41	122.4	3.9	72	16	AAV24970	Human gene signatu
42	60.2	3.2	65	24	ABN53183	Mouse spliced tran
43	60	3.2	60	24	ABN36212	Human spliced tran
44	45.4	2.5	40862	24	ABU34073	Human immune syste
45	43	2.3	469	23	ABV56638	Human prostate exp

#### ALIGNMENTS

RESULT 1  
AAV82779 standard; CDNA; 1851 BP.  
AAV82779;  
25-FEB-1999 (first entry)  
Clone bu45\_2 isolated from human adult placenta CDNA library.  
Secreted protein; nutritional activity; immune stimulating; vaccine;  
suppressing activity; haematopoiesis regulating activity;  
tissue growth activity; activin; inhibin activity; chemotaxis;  
chemokine activity; haemostasis; thrombolytic activity; receptor;  
ligand; anti-inflammatory; cadherin; tumour invasion suppressor;  
tumour inhibition; gene therapy; ds.  
Homo sapiens.  
WO9842739-A2.  
01-OCT-1998.  
20-MAR-1998; 98MO-US05653.  
19-MAR-1998; 98US-004446.  
21-MAR-1997; 97US-0822167.  
(GENY ) GENETICS INST INC.  
Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

PI Racie LA, Spaulding V, Treacy M;  
 XX MPI: 1998-609890/51.  
 DR P-PSDB; AAM85456.  
 XX  
 PT New polynucleotides encoding secreted human proteins - derived from  
 PT human foetal brain, adult brain, foetal kidney, placenta or adult  
 PT pineal gland cDNA libraries.  
 XX  
 PS Claim 14; Page 69-70; 113pp; English.  
 XX  
 CC The present sequence encodes a secreted protein. The polynucleotide and  
 CC secreted protein are predicted to have biological activities which would  
 CC make them suitable for treating, preventing or ameliorating medical  
 CC conditions in humans and animals, although no supporting data is given.  
 CC Suggested activities include nutritional activity, immune stimulating  
 CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating  
 CC activity, tissue growth activity, activin/inhibin activity,  
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,  
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour  
 CC invasion suppressor activity, and tumour inhibition activity (no data is  
 CC given in the specification to support these activities). The  
 CC polynucleotide is also stated to be useful for gene therapy.  
 CC  
 SQ Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;  
 Query Match 100.0%; Score 1851; DB 19; Length 1851;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1851; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTAGGCGCGAGCTTAGTCTCTGGAGCGCGCTCCGTGCGCGCGTCCGAGCCGCCCTTA 60  
 DB 1 GCGTAGGCGCGAGCTTAGTCTCTGGAGCGCGCTCCGTGCGCGCGTCCGAGCCGCCCTTA 60  
 QY 61 TCAGATATCTTAAACAGAAACCAACCTGGAAGAAAAAATGAATTCCTTATCTTGCGAT 120  
 DB 61 TCAGATATCTTAAACAGAAACCAACCTGGAAGAAAAAATGAATTCCTTATCTTGCGAT 120  
 QY 121 TTTTGGGAGGTGATACCTTTATCCGTGCTGCTGGGAAGCTATATGCAAGATGCA 180  
 DB 121 TTTTGGGAGGTGATACCTTTATCCGTGCTGCTGGGAAGCTATATGCAAGATGCA 180  
 QY 181 TCTTAAAGAGACTTTTGAAGAAATPAAAGAAATAGCCAGCTGGAGATGTTGCTA 240  
 DB 181 TCTTAAAGAGACTTTTGAAGAAATPAAAGAAATAGCCAGCTGGAGATGTTGCTA 240  
 QY 241 AACGATATCACTAAGCTGTTTATGTTAAGCCAGAACAGATCTATAGCGATTGG 300  
 DB 241 AACGATATCACTAAGCTGTTTATGTTAAGCCAGAACAGATCTATAGCGATTGG 300  
 QY 301 CACTCTGGTGTATCTGTTGAGCCAGACTGAGTCCCAAGAACCTTAAAGAAAGCA 360  
 DB 301 CACTCTGGTGTATCTGTTGAGCCAGACTGAGTCCCAAGAACCTTAAAGAAAGCA 360  
 QY 361 TCCAAATTAATGACAAAACCTGACAGAAAGATGGCTGGAGAAATTTCACTGGAGCAG 420  
 DB 361 TCCAAATTAATGACAAAACCTGACAGAAAGATGGCTGGAGAAATTTCACTGGAGCAG 420  
 QY 421 TGAGAAATACCCCACTGGAGAGAGGAGAAATCACTGTGTATGCTGAGCCAAAGATTC 480  
 DB 421 TGAGAAATACCCCACTGGAGAGAGGAGAAATCACTGTGTATGCTGAGCCAAAGATTC 480  
 QY 481 ATTAATATAGCCATCTGCGGTCTTGGCAGAGACTTGGACTCTCCAGAAAGCATTAACG 540  
 DB 481 ATTAATATAGCCATCTGCGGTCTTGGCAGAGACTTGGACTCTCCAGAAAGCATTAACG 540  
 QY 541 CAGAAGTCTGGTGTGAGACTCTTTCGATGAACTGACAGAGAGGSCCTCAGAGAGAGAG 600  
 DB 541 CAGAAGTCTGGTGTGAGACTCTTTCGATGAACTGACAGAGAGGSCCTCAGAGAGAGAG 600  
 QY 601 GGAAGATTTGTTTATTAACCAACCTTACATCACTCAAGAGCGGTGCAATACGGA 660  
 DB 601 GGAAGATTTGTTTATTAACCAACCTTACATCACTCAAGAGCGGTGCAATACGGA 660

QY 661 CGCAGGGGGGCGGTGGAAGCTGCGCAAGGTGGGGGCTTTGSCATCTCTCATTCGATCCGTGG 720  
 DB 661 CGCAGGGGGGCGGTGGAAGCTGCGCAAGGTGGGGGCTTTGSCATCTCTCATTCGATCCGTGG 720  
 QY 721 CCTCTTCTCCATCTACAGTCTCTCACAAGGTATTCAGAAATACAGAGATGGCGTCCCA 780  
 DB 721 CCTCTTCTCCATCTACAGTCTCTCACAAGGTATTCAGAAATACAGAGATGGCGTCCCA 780  
 QY 781 AAATTCACAGCCTGTATTCGCTGGAAGATGCAAAATGANTCAAGATGGCTTCTC 840  
 DB 781 AAATTCACAGCCTGTATTCGCTGGAAGATGCAAAATGANTCAAGATGGCTTCTC 840  
 QY 841 ATGGATCAAAATTTGTATTCAGTAAAGATGGGGGCAAGACCTTACCAGATCTGAT 900  
 DB 841 ATGGATCAAAATTTGTATTCAGTAAAGATGGGGGCAAGACCTTACCAGATCTGAT 900  
 QY 901 CCTTCAACACTGTATGACAGATCACTGGAGCAAAATTCAGAACAGTTGTACTGTCA 960  
 DB 901 CCTTCAACACTGTATGACAGATCACTGGAGCAAAATTCAGAACAGTTGTACTGTCA 960  
 QY 961 GTGACATCTGACAGCTGGAGTGTGGGCAAGGCTCCATGATGATGCGGTGAGCCT 1020  
 DB 961 GTGACATCTGACAGCTGGAGTGTGGGCAAGGCTCCATGATGATGCGGTGAGCCT 1020  
 QY 1021 TTATATCATGGGAAGCACTCTCATTTAAAGATCTTGGGCTGCGCCCAAGAGGAGCTC 1080  
 DB 1021 TTATATCATGGGAAGCACTCTCATTTAAAGATCTTGGGCTGCGCCCAAGAGGAGCTC 1080  
 QY 1081 TGCGGCTGTGCTCTGAGCTGACAGAAAGACAGGTGAGTGTGCTCCATGATATTC 1140  
 DB 1081 TGCGGCTGTGCTCTGAGCTGACAGAAAGACAGGTGAGTGTGCTCCATGATATTC 1140  
 QY 1141 AGTTACACAAAGTAATATTTCCACTACAGTCTGTGATGAGTCTGAGCGAGAACT 1200  
 DB 1141 AGTTACACAAAGTAATATTTCCACTACAGTCTGTGATGAGTCTGAGCGAGAACT 1200  
 QY 1201 TCTTACCACTGGGCTGCAATTCACCTGCGAGTAAAGGCGGAGGCGCATATGAGAGAG 1260  
 DB 1201 TCTTACCACTGGGCTGCAATTCACCTGCGAGTAAAGGCGGAGGCGCATATGAGAGAG 1260  
 QY 1261 TTATAGCCTGTGACAGCCCTCAATATCACTCAGTCTGTAGCCATGAGAGAGAGCAG 1320  
 DB 1261 TTATAGCCTGTGACAGCCCTCAATATCACTCAGTCTGTAGCCATGAGAGAGAGCAG 1320  
 QY 1321 ACATCAACTTTTGGATCCAAAGCTGAGTGGCTGAGGCACTTATGATGATTAACA 1380  
 DB 1321 ACATCAACTTTTGGATCCAAAGCTGAGTGGCTGAGGCACTTATGATGATTAACA 1380  
 QY 1381 AGTATTTCTTCTTCATCACTCCCAAGGAGACACATGATGATGATGATCAAAAGCAG 1440  
 DB 1381 AGTATTTCTTCTTCATCACTCCCAAGGAGACACATGATGATGATGATCAAAAGCAG 1440  
 QY 1441 TGAATGTGCTGTGCTGTGTTTGGGCTGTGTTCTTATGTTGTGACAGATGAGAGAA 1500  
 DB 1441 TGAATGTGCTGTGCTGTGTTTGGGCTGTGTTCTTATGTTGTGACAGATGAGAGAA 1500  
 QY 1501 TGTCTGCTAGGTCTTAAAGACATGAGAAAGAAAGCTTTTCACTGCTTGGCCAGGATC 1560  
 DB 1501 TGTCTGCTAGGTCTTAAAGACATGAGAAAGAAAGCTTTTCACTGCTTGGCCAGGATC 1560  
 QY 1561 CTGGGTCTGCAACTTTGAGAAACCTCTTCACTAATCAATTCATGCAATTCATCTCA 1620  
 DB 1561 CTGGGTCTGCAACTTTGAGAAACCTCTTCACTAATCAATTCATGCAATTCATCTCA 1620  
 QY 1621 AAGCAACTCTATTTATGCTTCTGTATTAATCTTCTTGATTAATCTTCAAAATTCCT 1680  
 DB 1621 AAGCAACTCTATTTATGCTTCTGTATTAATCTTCTTGATTAATCTTCAAAATTCCT 1680  
 QY 1681 GATTTGAAAAAAGAAATCAATTCCTCCCTCTCCCAACATGAAATCAATATGTA 1740  
 DB 1681 GATTTGAAAAAAGAAATCAATTCCTCCCTCTCCCAACATGAAATCAATATGTA 1740





```

Db      661 CGCAGGCGGCGGTGAGAGCTGCCAAGTGGGGCTTTGGCATCTCTCATTCAGTCCTGG 720
Qy      721 CTTCTCTTCACATCTACAGTCTCTCACAGATATTCAGGAATACCGAGATGGCGGCCCA 780
Db      721 CTTCTCTTCACATCTACAGTCTCTCACAGATATTCAGGAATACCGAGATGGCGGCCCA 780
Qy      781 AAATTCACAAGCCTGTATTAAGGAGATGACAGAAATGATGTCAGAAATGGCTTCTC 840
Db      781 AAATTCACAAGCCTGTATTAAGGAGATGACAGAAATGATGTCAGAAATGGCTTCTC 840
Qy      841 ATGGGATCAAAATTTGTCATTACGCTAAGAGTGGGGCAAGACTTACCCAGATCTGATT 900
Db      841 ATGGGATCAAAATTTGTCATTACGCTAAGAGTGGGGCAAGACTTACCCAGATCTGATT 900
Qy      901 CCTCAACACTGTAGAGAGATCACTGGGAGCAAAATACAGAAAGGTTGATCTGTCA 960
Db      901 CCTCAACACTGTAGAGAGATCACTGGGAGCAAAATACAGAAAGGTTGATCTGTCA 960
Qy      961 GTGACATCTGACAGCTGGAGATGTTGGCAGGGGTCCATGATGATGCGGTGGAGCCT 1020
Db      961 GTGACATCTGACAGCTGGAGATGTTGGCAGGGGTCCATGATGATGCGGTGGAGCCT 1020
Qy      1021 TTATATCATGGAAGACACTCTTATTAAGATCTTGGCTGCGTCCAAAGAGACTC 1080
Db      1021 TTATATCATGGAAGACACTCTTATTAAGATCTTGGCTGCGTCCAAAGAGACTC 1080
Qy      1081 TGGGCTGGTGTCTGAGACTGAGAAAGCAAGTGGAGTGGCTTCCAGTATATTC 1140
Db      1081 TGGGCTGGTGTCTGAGACTGAGAAAGCAAGTGGAGTGGCTTCCAGTATATTC 1140
Qy      1141 AGTTACACAAGGTAAATATTTTCCAACTACAGTCTGGTGAATGAGTCTGACGAGAACT 1200
Db      1141 AGTTACACAAGGTAAATATTTTCCAACTACAGTCTGGTGAATGAGTCTGACGAGAACT 1200
Qy      1201 TCTTACCCACTGGGCTGCAATTCACCTGCGAGTGAAGAGCCAGTCACTGAGAGAG 1260
Db      1201 TCTTACCCACTGGGCTGCAATTCACCTGCGAGTGAAGAGCCAGTCACTGAGAGAG 1260
Qy      1261 TTATAGGCTGCTGAGAGCCCTCAATATGATCAGTCTGAGCCATGAGAGAGAGAG 1320
Db      1261 TTATAGGCTGCTGAGAGCCCTCAATATGATCAGTCTGAGCCATGAGAGAGAGAG 1320
Qy      1321 ACATCAACTTTTGAATCCAGAGTGGAGTCCCTGAGCCAGTCTGATGATGATTTACA 1380
Db      1321 ACATCAACTTTTGAATCCAGAGTGGAGTCCCTGAGCCAGTCTGATGATGATTTACA 1380
Qy      1381 AGTATTTCTTCTTCATCACTCCCAAGAGACCAATGATGATGATCCAAAGCAGA 1440
Db      1381 AGTATTTCTTCTTCATCACTCCCAAGAGACCAATGATGATGATCCAAAGCAGA 1440
Qy      1441 TGAATGTTGCTGCTGTTGGGCTGTTTCTTATGTTGTCAGACATGAGAGAA 1500
Db      1441 TGAATGTTGCTGCTGTTGGGCTGTTTCTTATGTTGTCAGACATGAGAGAA 1500
Qy      1501 TGTCTGCTAGGTCTCTAGAAACAGTAAGAAAGAGTTCATGCTTGGCAGAGATC 1560
Db      1501 TGTCTGCTAGGTCTCTAGAAACAGTAAGAAAGAGTTCATGCTTGGCAGAGATC 1560
Qy      1561 CTGGGCTGCAACTTTGGAAAATCCTCTTCAATPAACAATTCATTCATCTTCA 1620
Db      1561 CTGGGCTGCAACTTTGGAAAATCCTCTTCAATPAACAATTCATTCATCTTCA 1620
Qy      1621 AAGCAACTCTATTTCAATGCTTCTGTTATATCTTCTTGAATCTTCCAAATTCCT 1680
Db      1621 AAGCAACTCTATTTCAATGCTTCTGTTATATCTTCTTGAATCTTCCAAATTCCT 1680
Qy      1681 GATTCTAGAAAAGAAATCAATTCCTCCCTCCCTCCCAACAATCAATCAATGATG 1740
Db      1681 GATTCTAGAAAAGAAATCAATTCCTCCCTCCCTCCCAACAATCAATCAATGATG 1740
Qy      1741 GGGATTAACAGTGGGGCAATTTCTTATATCACTCTTAAACATGTTTCCACTTTAA 1800
Db      1741 GGGATTAACAGTGGGGCAATTTCTTATATCACTCTTAAACATGTTTCCACTTTAA 1800

```

```

Qy      1801 AGTAACACTTAATTAATTTTGAAGATCTCTGAAAAAAGAAAAA 1851
Db      1801 AGTAACACTTAATTAATTTTGAAGATCTCTGAAAAAAGAAAAA 1851

RESULT 3
AA298034
ID AA298034 standard; cDNA, 1863 BP.
XX
AC AA298034;
XX
DT 09-MAY-2000 (first entry)
XX
DE Human secreted protein encoding nucleotide sequence SEQ ID NO:28.
XX
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW antiinflammatory; neurotropic; neuroprotective; antiallergic; cancer;
KW tumour; neurodegenerative disorder; developmental abnormality; allergy;
KW foetal deficiency; blood disorder; immune system disorder; arthritis;
KW autoimmune disease; hepatic disease; renal disease; inflammation;
KW Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
KW infection; AIDS; spinal cord injury; transplant rejection; diabetes;
KW asthma; sepsis; acne; psoriasis; cardiovascular disorder;
KW reproductive disorder; gastrointestinal disorder; respiratory disorder;
KW metabolic disorder; food additive; preservative; ss.

OS Homo sapiens.
XX
PN WO200004140-A1.
XX
PD 27-JAN-2000.
XX
PF 14-JUL-1999; 99WO-US15849.
XX
PR 15-JUL-1998; 98US-0092921.
PR 15-JUL-1998; 98US-0092922.
PR 15-JUL-1998; 98US-0092956.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
PI Lafleur DW, Ehner R, Olsen HS, Brewer LA, Florence KA, Young PE;
PI Muceneki M, Endress GA, Soppet DR;
XX
XX WPI: 2000-161128/14.
DR P-PSDB; AAY87081.
XX
PT New isolated human genes, useful for diagnosis and treatment of, e.g.
PT cancers, neurological or blood disorders
XX
PS Claim 1; Page 319; 494pp; English.
XX

The polynucleotide sequences given in AA298017 to AA298108 encode the
CC human secreted proteins given in AAY87064 to AAY87223. Human secreted
CC protein can have activities based on the tissues and cells the genes are
CC expressed in. Examples of activities include: cytostatic;
CC immunosuppressive; antiinflammatory; neurotropic; neuroprotective; and
CC antiallergic. The polynucleotides and their corresponding secreted
CC polypeptides are useful for preventing, treating or ameliorating medical
CC conditions, e.g. by protein or gene therapy. Also pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the new
CC polynucleotides. Human secreted protein s and their polynucleotides can
CC be used for developing products for the diagnosis or treatment of cancer,
CC tumours, neurodegenerative disorders, developmental abnormalities and
CC foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic disorders, diseases of the immune system,
CC allergies, Alzheimer's disease, behavioural disorders, schizophrenia,
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
CC cardiovascular disorders, reproductive disorders, gastrointestinal
CC disorders, respiratory disorders and metabolic disorders. The

```

CC proteins or polynucleotides can also be used as food additives or  
CC preservatives. The proteins are also useful for identifying their  
CC binding partners. AA298008 to AA298016 and AA298063 are sequence used in  
CC the exemplification of the present invention.

XX Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Query Match 99.7%; Score 1846.2; DB 21; Length 1863;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1846; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 GGCTAGGCCGCGAGCTTATGCTTGGAGCGCGCTTCCGCGCGCTGAGAGCGCCCTTA 60
DB 1 GACTAGGCCGCGAGCTTATGCTTGGAGCGCGCTTCCGCGCGCTGAGAGCGCCCTTA 60
QY 61 TCAGATTTATCTTAAACAAGAAACCAACCTGGAAGAAATGAAATTCCTTATCTTCCAT 120
DB 61 TCAGATTTATCTTAAACAAGAAACCAACCTGGAAGAAATGAAATTCCTTATCTTCCAT 120
QY 121 TTTTCGGTGGTTCACCTTTTATCCCTGCTCTGGGAAGCTATATGCAAGATGCA 180
DB 121 TTTTCGGTGGTTCACCTTTTATCCCTGCTCTGGGAAGCTATATGCAAGATGCA 180
QY 181 TCTCTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTGCTA 240
DB 181 TCTCTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTGCTA 240
QY 241 AAGCAATCATCAACCTAGCTGTTTATGTGTAAGCCCAAGACATCTATAGGATGG 300
DB 241 AAGCAATCATCAACCTAGCTGTTTATGTGTAAGCCCAAGACATCTATAGGATGG 300
QY 301 CACTTCTGTGATACGTGTTGAGCCCAAGCTGAGTGGCTCCAGAACTTAGAAAGCA 360
DB 301 CACTTCTGTGATACGTGTTGAGCCCAAGCTGAGTGGCTCCAGAACTTAGAAAGCA 360
QY 361 TCCAAATTAATGATACCAAACTGACAGCAAGATGGGCTGGAAGAAATTCACCTGAGCCAG 420
DB 361 TCCAAATTAATGATACCAAACTGACAGCAAGATGGGCTGGAAGAAATTCACCTGAGCCAG 420
QY 421 TGAAGATACCCCACTGGAGAGGGAGAAATACAGCTGTGATGTGTGAGCCAGAAATTC 480
DB 421 TGAAGATACCCCACTGGAGAGGGAGAAATACAGCTGTGATGTGTGAGCCAGAAATTC 480
QY 481 ATAAGATAGCCATCTGGGCTGTGGCAGCAGATGGGACTCTTCAGAAAGCAATTAAG 540
DB 481 ATAAGATAGCCATCTGGGCTGTGGCAGCAGATGGGACTCTTCAGAAAGCAATTAAG 540
QY 541 CAGAAGTCTGTGTGACTCTTTCGATGAACTGCAAGAGAGGGCTCAGAAAGCAAGAG 600
DB 541 CAGAAGTCTGTGTGACTCTTTCGATGAACTGCAAGAGAGGGCTCAGAAAGCAAGAG 600
QY 601 GGAAGATGTTGTTTATTAACAACCTTAACATCACTCAAGGACGGTGAATACCGAA 660
DB 601 GGAAGATGTTGTTTATTAACAACCTTAACATCACTCAAGGACGGTGAATACCGAA 660
QY 661 CGCAGGGGGCGGTGGAAGCTGCAAGGTTGGGGCTTTGGCATCTCTCAATGATCCGGG 720
DB 661 CGCAGGGGGCGGTGGAAGCTGCAAGGTTGGGGCTTTGGCATCTCTCAATGATCCGGG 720
QY 721 CCTCTCTCCATCTACAGTCTCTCAACAGGATATCAGGAATACAGAGATGGCTGCCA 780
DB 721 CCTCTCTCCATCTACAGTCTCTCAACAGGATATCAGGAATACAGAGATGGCTGCCA 780
QY 781 AAATTCACAAGCTGTATTAAGGTGAAGATGCAAGAAATGATCAAGATGCTTCTC 840
DB 781 AAATTCACAAGCTGTATTAAGGTGAAGATGCAAGAAATGATCAAGATGCTTCTC 840
QY 841 ATGGATCAAAATGTTCTACGCTAAGATGAGGGGCAAGACCTAATCCAGATCTGATT 900
DB 841 ATGGATCAAAATGTTCTACGCTAAGATGAGGGGCAAGACCTAATCCAGATCTGATT 900
QY 901 CCTTCACAAGCTGTAGAGATCACTGGAGCAAAATATCCAGAAAGGTTGACTGTCA 960

```

```

DB 901 CCTTCACAAGCTGTAGAGATCACTGGAGCAAAATATCCAGAAAGGTTGACTGTCA 960
QY 961 GTGACATCTGACAGCTGGAGATGTTGGCAGGGGTGCATGATGATGGCGTGAAGCT 1020
DB 961 GTGACATCTGACAGCTGGAGATGTTGGCAGGGGTGCATGATGATGGCGTGAAGCT 1020
QY 1021 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGGAGCTC 1080
DB 1021 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGGAGCTC 1080
QY 1081 TGGGCTGTGCTGTGACATGCAAGAGAACAGGTGAGATGTTGGCTTCCAGTATATC 1140
DB 1081 TGGGCTGTGCTGTGACATGCAAGAGAACAGGTGAGATGTTGGCTTCCAGTATATC 1140
QY 1141 AGTTACACAAGGTAATATTTTCCAACTACAGCTGCTGATGATGAGTCAAGCGAGAACT 1200
DB 1141 AGTTACACAAGGTAATATTTTCCAACTACAGCTGCTGATGATGAGTCAAGCGAGAACT 1200
QY 1201 TCTTACCACTGGGCTGCAATTCACCTGCAAGTGAAGAAAGCCAGGCGCATATGAGAGAG 1260
DB 1201 TCTTACCACTGGGCTGCAATTCACCTGCAAGTGAAGAAAGCCAGGCGCATATGAGAGAG 1260
QY 1261 TTATGACCTGTGACAGCCCTTCAATATCACTAGTCTTGAAGCATGAGAGAGAGAG 1320
DB 1261 TTATGACCTGTGACAGCCCTTCAATATCACTAGTCTTGAAGCATGAGAGAGAGAG 1320
QY 1321 ACATCAACTTTTGGATTCAGAGCTGAGATGCTGAGAGCACTTATGATGACTATACA 1380
DB 1321 ACATCAACTTTTGGATTCAGAGCTGAGATGCTGAGAGCACTTATGATGACTATACA 1380
QY 1381 AGTATTTCTTCTTCATCACTCCCAAGAGACACATGATGATGATGATGATGATGATG 1440
DB 1381 AGTATTTCTTCTTCATCACTCCCAAGAGACACATGATGATGATGATGATGATGATG 1440
QY 1441 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTTGACAGATGGAAGAA 1500
DB 1441 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTTGACAGATGGAAGAA 1500
QY 1501 TGTGCTAGGCTCTTAAACAGTAAGAAAGAAAGCTTTCATGCTTCTGCGCAGGAATC 1560
DB 1501 TGTGCTAGGCTCTTAAACAGTAAGAAAGAAAGCTTTCATGCTTCTGCGCAGGAATC 1560
QY 1561 CTGGGCTGCAACTTTGGAAGAACTCTCTTCAACATTAACATTTATCAATTCATCTTCA 1620
DB 1561 CTGGGCTGCAACTTTGGAAGAACTCTCTTCAACATTAACATTTATCAATTCATCTTCA 1620
QY 1621 AAGCAAACTTATTCATGCTTCTGTTATATCTTCTTATGATGATCTTCAAAATCTCT 1680
DB 1621 AAGCAAACTTATTCATGCTTCTGTTATATCTTCTTATGATGATCTTCAAAATCTCT 1680
QY 1681 GATTCGTGAAGAAAGGATCATCTTCCCTGCTCCCAACATGATGATGATGATGATG 1740
DB 1681 GATTCGTGAAGAAAGGATCATCTTCCCTGCTCCCAACATGATGATGATGATGATG 1740
QY 1741 GGAATTCAGAGGGGAGATTTCTTATATCACTCTTAAAGAAAGCTGTTTCACTTTAA 1800
DB 1741 GGAATTCAGAGGGGAGATTTCTTATATCACTCTTAAAGAAAGCTGTTTCACTTTAA 1800
QY 1801 AGTAAACACTTAATTAATTTTGGAGATCTCTGAAAAAAGAAAAAAGAAAAA 1851
DB 1801 AGTAAACACTTAATTAATTTTGGAGATCTCTGAAAAAAGAAAAAAGAAAAA 1851

```

RESULT 4  
AD11647  
ID AD11647 standard; cDNA, 1863 BP.  
XX AD11647;  
XX AC  
XX 24-SEP-2001 (first entry)  
DE Human secreted protein-encoding gene 18 cDNA clone HRAU35, SEQ ID NO:28.

KM	Human; secreted protein; proliferative disorder; cancer; tumour; ashma;
KW	foetal abnormality; developmental abnormality; haematopoietic disorder;
KW	immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KV	parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
KW	psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
KW	inflammation; neurological disorder; Alzheimer's disease; food additive;
KM	angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
KM	pregnancy-related disorder; endocrine disorder; infection; wound healing;
KW	cell culture; chemotaxis; vulnerability; binding partner identification;
KW	gene therapy; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS 99..1517
FT	/tag= a
FT	/product= "Human secreted protein precursor"
FT	sig_peptide 99..170
FT	/tag= b
FT	mat_peptide 171..1514
FT	/tag= c
FT	/product= "Mature human secreted protein"
XX	
PM	WO200151504-A1.
XX	
PD	19-JUL-2001.
XX	
PF	12-JAN-2001; 2001WO-US00911.
XX	
PR	13-JAN-2000; 2000US-0482273.
XX	
PA	(HDMA-) HUMAN GENOME SCI INC.
XX	
P1	Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y;
P1	Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;
P1	Endress GA, Muscinski M, Ebner R;
XX	
XK	WPI; 2001-425865/45.
DR	P-PSDB; AAE06058.
XX	
PT	Isolated nucleic acid molecule encoding a human secreted protein is
PT	used in preventing, treating or ameliorating a medical condition -
XX	
PS	Claim 1; Page 686; 864pp; English.
XX	
CC	AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted
CC	protein genes, and AAE06041-AAE06132 represent the proteins they encode.
CC	AAE06133-AAE06205 represent human secreted protein fragments.
CC	The secreted proteins and their genes are useful for preventing, treating,
CC	or ameliorating medical conditions, e.g., by protein or gene therapy.
CC	Pathological conditions can be diagnosed by determining the amount of the
CC	new protein in a sample or by determining the presence of mutations in
CC	the new genes. Specific uses are described for each of the 71 genes,
CC	based on the tissues in which they are most highly expressed, and include
CC	developing products for the diagnosis or treatment of proliferative
CC	disorders, cancer, tumours, foetal and developmental abnormalities,
CC	haematopoietic disorders, diseases of the immune system, AIDS, autoimmuna-
CC	diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC	neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC	cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC	psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC	angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC	pregnancy-related disorders, endocrine disorders, and infections. The
CC	proteins can also be used to aid wound healing and epithelial cell
CC	proliferation, to prevent skin aging due to sunburn, to maintain organs
CC	before transplantation, for supporting cell culture of primary tissues,
CC	to regenerate tissues, to identify their cognate ligands or binding
CC	partners, and in chemotaxis, and can be used as a food additive or
CC	preservative to modify storage properties. Antibodies specific for a
CC	protein of the invention can be used in alleviating symptoms associated
CC	with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC	radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
CC	The present sequence represents a human secreted protein-encoding cDNA of

[illegible]

```

Db      961 GTGACATCTGACAGCTGGAGTGTGGGAGGGTGCATGATGAGCGGTGAGACCT 1020
Qy      1021 TTATATATGAGGAGAGCACTCTCACTTATTAAGATCTTGGCTGCGCTCCAAAGGAGCTC 1080
Db      1021 TTATATATGAGGAGAGCACTCTCACTTATTAAGATCTTGGCTGCGCTCCAAAGGAGCTC 1080
Qy      1081 TGGCGCTGCTGCTCTGAGCTGAGAGAGACAGAGGTGAGTGGCTTCCAGATTAATC 1140
Db      1081 TGGCGCTGCTGCTCTGAGCTGAGAGAGACAGAGGTGAGTGGCTTCCAGATTAATC 1140
Qy      1141 AGTTACACAAAGTAAATATTTTCCAACTACAGTCTGTGATGAGTCTGACGACAGAACT 1200
Db      1141 AGTTACACAAAGTAAATATTTTCCAACTACAGTCTGTGATGAGTCTGACGACAGAACT 1200
Qy      1201 TCTTACCCACTGGGCTGCAATTCATCTGAGAGTGAAGGCGGCGCATCATGAGAGAG 1260
Db      1201 TCTTACCCACTGGGCTGCAATTCATCTGAGAGTGAAGGCGGCGCATCATGAGAGAG 1260
Qy      1261 TTATGAGCCTGCTGCTGAGCCCTCAATATCACTCAGTCTGTGAGCCATGAGAGAGAG 1320
Db      1261 TTATGAGCCTGCTGCTGAGCCCTCAATATCACTCAGTCTGTGAGCCATGAGAGAGAG 1320
Qy      1321 ACATCAACTTTTGGATCCAGCTGAGAGTGGAGCCAGTCTAATTGATGACTTATACA 1380
Db      1321 ACATCAACTTTTGGATCCAGCTGAGAGTGGAGCCAGTCTAATTGATGACTTATACA 1380
Qy      1381 AGATATTTCTTCTTCCATCATCTCCACAGAGACACCATGATGATGATGATGATGATG 1440
Db      1381 AGATATTTCTTCTTCCATCATCTCCACAGAGACACCATGATGATGATGATGATGATG 1440
Qy      1441 TGAATGTTGCTGCTGCTGTTTGGGCTGTTTCTTATGTTGTTGTCAGACATGAGAGAA 1500
Db      1441 TGAATGTTGCTGCTGCTGTTTGGGCTGTTTCTTATGTTGTTGTCAGACATGAGAGAA 1500
Qy      1501 TGTGCTGCTGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1560
Db      1501 TGTGCTGCTGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1560
Qy      1561 CTGGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db      1561 CTGGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Qy      1621 AAGCACAACCTATTTTATGAGTCTGTATATCTTATGATCTTATGATCTTATGATCT 1680
Db      1621 AAGCACAACCTATTTTATGAGTCTGTATATCTTATGATCTTATGATCTTATGATCT 1680
Qy      1681 GATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db      1681 GATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Qy      1741 GGGATTACAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db      1741 GGGATTACAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Qy      1801 AGTAAACACTTAATTAATTTTGAAGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAG 1851
Db      1801 AGTAAACACTTAATTAATTTTGAAGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAG 1851

```

```

KW      virtucide; fungicide; ophthalmological; autoimmune disease; neoplasm;
KW      rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
KW      cardiovascular disorder; cerebrovascular disorder; cerebral ischemia;
KW      angiodysplasia; nervous system disorder; Alzheimer's disease; infection;
KW      ocular disorder; corneal infection; wound healing; skin aging;
KW      epithelial cell proliferation; food additive.
XX
OS      Homo sapiens.
XX
PN      NC0200226931-A2.
XX
PD      04-APR-2002.
XX
PF      24-SEP-2001; 2001MO-US29871.
XX
PR      25-SEP-2000; 2000US-234925P.
PR      12-JAN-2001; 2001MO-US00911.
XX
PA      (HUMA-) HUMAN GENOME SCI INC.
PI      Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y;
PI      Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;
PI      Endress GA, Mucenski M, Ebner R;
XX
DR      WPI; 2002-362489/39.
DR      P-PSDB; ABG33880.
XX
PT      Novel 71 isolated secreted polypeptides and polynucleotides encoding
PT      the polypeptides, useful for treating Huntington's disease, sepsis,
PT      meningitis, thrombocytopenia, haemolytic anaemia, rheumatoid arthritis,
PT      asthma
XX
PS      Claim 1; Page 1170; 1478pp; English.
XX
CC      The invention relates to an isolated nucleic acid molecule (or its
CC      fragment, homologue complement or allelic variant) encoding a human
CC      secreted protein (and its fragment, domain, epitope, variant, secreted
CC      form and species variant). Also included are a recombinant vector
CC      comprising the nucleic acid, a recombinant host cell comprising the
CC      vector, an antibody against the secreted protein, a recombinant host cell
CC      that expresses the secreted protein and a method of identifying a binding
CC      partner of the secreted protein. The nucleic acid and protein are used to
CC      prevent, diagnose, treat or ameliorate a medical condition in e.g.
CC      humans, mice, rabbits, goats, cats, dogs, chickens or sheep
CC      for example autoimmune diseases e.g. rheumatoid arthritis,
CC      hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC      cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC      e.g. cerebral ischemia, angiodysplasia, nervous system disorders e.g.
CC      Alzheimer's disease, infections caused by bacteria, viruses and fungi and
CC      ocular disorders e.g. corneal infection. Many other diseases and
CC      disorders are listed in the specification. The polypeptides can also be
CC      used to aid wound healing an epithelial cell proliferation, to prevent
CC      skin aging due to sunburn, to maintain organs before transplantation, for
CC      supporting cell culture of primary tissues, to regenerate tissues and in
CC      chemotaxis. The polypeptides can also be used as a food additive or
CC      preservative to increase or decrease storage capabilities. The present
CC      sequence encodes a novel human secreted protein of the invention.
XX
SQ      Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

```

Query Match 99.7%; Score 1846.2; DB 24; Length 1863;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1846; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy      1 GGCTAGGCGCGAGCTTATGCTGGAGCGGCTCGCTGCGCGGCTGAGAGCGGCGCTA 60
Db      1 GACTAGGCGCGAGCTTATGCTGGAGCGGCTCGCTGCGCGGCTGAGAGCGGCGCTA 60
Qy      61 TCAGATTATCTTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db      61 TCAGATTATCTTACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Qy      121 TTTTCGGTGTGTTACCTTTATCCCTGTGCTGGAGAGAGATATATGACAGAGATGGA 180

```

```

Db 121 TTTTCGGTGTGTCACTTTATCCCTGTGCTGTGGAAAGCTATATGCAAGATGGA 180
Qy 181 TCTCTAAGGAGCTTTTGAAGAAATPAAAGAAATAGCCAGCTGTGAGATGTTGCTA 240
Db 181 TCTCTAAGGAGCTTTTGAAGAAATPAAAGAAATAGCCAGCTGTGAGATGTTGCTA 240
Qy 241 AAGCATCATCAACCTAGCTGTTTATGTGTAAGCCAGAAAGATCTTATGAGCATGG 300
Db 241 AAGCATCATCAACCTAGCTGTTTATGTGTAAGCCAGAAAGATCTTATGAGCATGG 300
Qy 301 CACTTCTGGTGTATCTGTTGAGCCAGACTGAGTGGCTCAAGAACTTGAAGAACCA 360
Db 301 CACTTCTGGTGTATCTGTTGAGCCAGACTGAGTGGCTCAAGAACTTGAAGAACCA 360
Qy 361 TCCAAATTAATGACAAACCTGACGAAGATGGCTGAGAAAGTTCACTGGAGCCAG 420
Db 361 TCCAAATTAATGACAAACCTGACGAAGATGGCTGAGAAAGTTCACTGGAGCCAG 420
Qy 421 TGAGAAATACCCCACTGGAGAGGGAGAAAGATCAGCTGTGATGCTGAGCCAGAAATTC 480
Db 421 TGAGAAATACCCCACTGGAGAGGGAGAAAGATCAGCTGTGATGCTGAGCCAGAAATTC 480
Qy 481 ATAGATAGCCATCTGAGGTCTGTGGCAGCAGCATTGGGACTCTCCAGAAAGCATTACAG 540
Db 481 ATAGATAGCCATCTGAGGTCTGTGGCAGCAGCATTGGGACTCTCCAGAAAGCATTACAG 540
Qy 541 CAGAAATCTGTGTGTGATCTCTTTCGATGAACTGACAGAAAGGCTTCAAGAAAGAG 600
Db 541 CAGAAATCTGTGTGTGATCTCTTTCGATGAACTGACAGAAAGGCTTCAAGAAAGAG 600
Qy 601 GGAAGATGTTGTTTATTAACAACCTTACATCACTCAAGAGAGGCTGATACCGAA 660
Db 601 GGAAGATGTTGTTTATTAACAACCTTACATCACTCAAGAGAGGCTGATACCGAA 660
Qy 661 CGCAGAGGCGGTGAGAGCTGCAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGG 720
Db 661 CGCAGAGGCGGTGAGAGCTGCAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGG 720
Qy 721 CCTCTCTTCCATCTACAGTCTCTCAACAGGTATTCAGGAATACAGAGATGGCGCCA 780
Db 721 CCTCTCTTCCATCTACAGTCTCTCAACAGGTATTCAGGAATACAGAGATGGCGCCA 780
Qy 781 AAATTCCAACAGCTGTATTAAGGTGAGAAAGATGCAAGAAATGATGCTCAAGATGCTTC 840
Db 781 AAATTCCAACAGCTGTATTAAGGTGAGAAAGATGCAAGAAATGATGCTCAAGATGCTTC 840
Qy 841 ATGGATCAAAATTTGTCAATTCAGCTAAAGATGGGGCAAAAGACTACCCAGATCTGAT 900
Db 841 ATGGATCAAAATTTGTCAATTCAGCTAAAGATGGGGCAAAAGACTACCCAGATCTGAT 900
Qy 901 CCTTCAACAGCTGTACAGATCACTGGAGCAAAATTCAGAAACAGGTGTGATCGTCA 960
Db 901 CCTTCAACAGCTGTACAGATCACTGGAGCAAAATTCAGAAACAGGTGTGATCGTCA 960
Qy 961 GTGACATCTGAGACAGTGGAGTGTGGCAGAGGTGCCATGATGAGCGGTGGAGCCT 1020
Db 961 GTGACATCTGAGACAGTGGAGTGTGGCAGAGGTGCCATGATGAGCGGTGGAGCCT 1020
Qy 1021 TTAATATCATGGAGACACTCTCACTTATTAAGATCTTGGGCTGGCTCAAGAGAGACTC 1080
Db 1021 TTAATATCATGGAGACACTCTCACTTATTAAGATCTTGGGCTGGCTCAAGAGAGACTC 1080
Qy 1081 TGGCGCTGTGTCTGTGAGCTGCAAGAAAGAAAGAGTGTGCTTCCAGTATTAATC 1140
Db 1081 TGGCGCTGTGTCTGTGAGCTGCAAGAAAGAAAGAGTGTGCTTCCAGTATTAATC 1140
Qy 1141 AGTTACACAAGGTAAATATTTCCAACTACAGCTGTGATGAGGTCTGACGAGAACTT 1200
Db 1141 AGTTACACAAGGTAAATATTTCCAACTACAGCTGTGATGAGGTCTGACGAGAACTT 1200
Qy 1201 TCTTACCACTGTGGCTGCAATTCATGCGAGTGAAGAGCCAGGGCATCATGAGAGAG 1260

```

```

Db 1201 TCTTACCACTGTGGCTGCAATTCATGCGAGTGAAGAGCCAGGGCCATCATGAGAGAG 1260
Qy 1261 TTAATAGCTGTGCTCAGCCCTCAATATCACTCAGTCTCGAGCCATGAGAGAGAGAG 1320
Db 1261 TTAATAGCTGTGCTCAGCCCTCAATATCACTCAGTCTCGAGCCATGAGAGAGAGAG 1320
Qy 1321 ACATCACTTTTGGATCCAGAGCTGAGAGTCCCTGAGAGCAGTCTTATGATGATTAACA 1380
Db 1321 ACATCACTTTTGGATCCAGAGCTGAGAGTCCCTGAGAGCAGTCTTATGATGATTAACA 1380
Qy 1381 AGTATTTCTTCTTCATCACTCCCAAGAGACACCATGATGATGATGATCAAGAGAGA 1440
Db 1381 AGTATTTCTTCTTCATCACTCCCAAGAGACACCATGATGATGATGATCAAGAGAGA 1440
Qy 1441 TGAATGTTGCTGCTGCTGTTGGGCTGTGTTCTTATGTTGTTGACAGATGAGAGAAA 1500
Db 1441 TGAATGTTGCTGCTGCTGTTGGGCTGTGTTCTTATGTTGTTGACAGATGAGAGAAA 1500
Qy 1501 TGTGCTGAGTCTCTAGAAACAGTAAAGAAAGAAAGCTTTTCATGCTTCTGGCCAGAAATC 1560
Db 1501 TGTGCTGAGTCTCTAGAAACAGTAAAGAAAGAAAGCTTTTCATGCTTCTGGCCAGAAATC 1560
Qy 1561 CTGGGCTGCAACTTTGGAAAACTCTCTTCACTAATCAATTCATTCATTCATTCATTC 1620
Db 1561 CTGGGCTGCAACTTTGGAAAACTCTCTTCACTAATCAATTCATTCATTCATTCATTC 1620
Qy 1621 AAGCAACATCTATTTTATGCTTCTGTTATATCTTCTTATGATCTTCAATTCCT 1680
Db 1621 AAGCAACATCTATTTTATGCTTCTGTTATATCTTCTTATGATCTTCAATTCCT 1680
Qy 1681 GATTCTGAAAAAAGAAATCATCTTCCCTCCCTCCCAACATGATGATGATGATGATGAT 1740
Db 1681 GATTCTGAAAAAAGAAATCATCTTCCCTCCCTCCCAACATGATGATGATGATGATGAT 1740
Qy 1741 GGGATTAAGTGGGGGATTTCTTATATCACTCTTAAAAACATGTTTCCACTTTAAA 1800
Db 1741 GGGATTAAGTGGGGGATTTCTTATATCACTCTTAAAAACATGTTTCCACTTTAAA 1800
Qy 1801 AGTAAACATTAATTAATTTTGAAGATCTCTGAAAAAAGAAAAA 1851
Db 1801 AGTAAACATTAATTAATTTTGAAGATCTCTGAAAAAAGAAAAA 1851

RESULT 6
ACC50817
ID ACC50817 standard; cDNA; 1863 BP.
XX
AC ACC50817;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human secreted protein coding sequence, SEQ ID 484.
XX
KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cyostatic;
XX vulnery; antiinflammatory; nootropic; neuroprotective;
XX antiparkinsonian; gene therapy; human; cardiovascular disorder;
XX gene; ss.
XX
OS Homo sapiens.
XX
PN MO200295010-A2.
XX
PD 28-NOV-2002.
XX
PF 19-MAR-2002; 2002WO-US09785.
XX
PR 21-MAR-2001; 2001US-27340P.
XX 19-JUN-2001; 2001US-306171P.
XX 13-NOV-2001; 2001US-331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;

```

XX WPI; 2003-129429/12.  
DR Novel human secreted proteins, useful for detecting, preventing,  
XX diagnosing, prognosticating, treating and/or ameliorating  
PT cardiovascular disorders such as arrhythmia -  
XX  
XX Claim 21; SEQ ID 484; 1881bp; English.  
XX  
CC The present invention relates to novel human secreted proteins  
CC (ABR47633-ABR48145) and their coding sequences (ACC50344-ACC50856). The  
CC proteins and their coding sequences are useful for the preparation of a  
CC diagnostic or pharmaceutical composition for diagnosing or treating a  
CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,  
CC coronary arteriosclerosis and myocardial ischemia), neural disorders,  
CC immune system disorders, muscular disorders, reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, renal disorders,  
CC proliferative disorders and/or cancerous diseases and conditions, for  
CC wound healing and epithelial cell proliferation, to treat inflammation or  
CC infection, for treating thrombosis and arteriosclerosis, for treating or  
CC preventing neural damage which occurs in neuronal disorders or  
CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
CC disease, to enhance bone and periodontal regeneration and aid in tissue  
CC transplants or bone grafts, to prevent skin aging or hair loss, to  
CC stimulate growth and differentiation of haematopoietic cells and bone  
CC marrow cells when used in combination with other cytokines, to maintain  
CC organs before transplantation or for supporting cell culture of primary  
CC tissues, to increase or decrease differentiation or proliferation of  
CC embryonic stem cells, or to modulate mammalian characteristics or  
CC metabolism.  
CC Note: The sequence data for this patent was published in electronic  
CC format and is available from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
CC  
XX Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;  
XX  
Query Match 99.7%; Score 1846.2; DB 25; Length 1863;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1846; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 GGCTAGGCGCGGAGCTTATGCTGGAGCCGCTCCGCTGCGCGCTCAGAGCCGCCCTA 60  
DB 1 GACTAGGCGCGGAGCTTATGCTGGAGCCGCTCCGCTGCGCGCTCAGAGCCGCCCTA 60  
QY 61 TCAGATTATCTTAAACAAGAAACCACTGGAAGAAATGAATTCCTTATCTTCCGAT 120  
DB 61 TCAGATTATCTTAAACAAGAAACCACTGGAAGAAATGAATTCCTTATCTTCCGAT 120  
QY 121 TTTTGGGCGGTCACTTTATCCCTGCTGCTGGGAAGCTATATGCAAGAAATGCA 180  
DB 121 TTTTGGGCGGTCACTTTATCCCTGCTGCTGGGAAGCTATATGCAAGAAATGCA 180  
QY 181 TCTCTAAGAGAGCTTTGAGAAATTAAGAAATAGCAGCTGTGAGATTTGCTA 240  
DB 181 TCTCTAAGAGAGCTTTGAGAAATTAAGAAATAGCAGCTGTGAGATTTGCTA 240  
QY 241 AAGCAATCATCAACCTAGCTGTTATGTTAAAGCCAGCAAGATCTATAGAGATTTG 300  
DB 241 AAGCAATCATCAACCTAGCTGTTATGTTAAAGCCAGCAAGATCTATAGAGATTTG 300  
QY 301 CACTTCGTTGATACGTGTGAGCCAGCTGAGTGGCTCAAGAACTAGAAAAGCA 360  
DB 301 CACTTCGTTGATACGTGTGAGCCAGCTGAGTGGCTCAAGAACTAGAAAAGCA 360  
QY 361 TCCAAATTAATGACCAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGGAGCCAG 420  
DB 361 TCCAAATTAATGACCAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGGAGCCAG 420  
QY 421 TGAGATATCCCACTGGAGAGGGAGAAAGATCACTGCTGTGATGCTGAGCCAGAAATTC 480  
DB 421 TGAGATATCCCACTGGAGAGGGAGAAAGATCACTGCTGTGATGCTGAGCCAGAAATTC 480  
QY 481 ATAAGATAGCATCTGGGTCTTGGCAGACAGATTGGAGCTCTCCAGAAAGCATTAAG 540

DB 481 ATAAGATAGCATCTGGGTCTTGGCAGACAGATTGGAGCTCTCCAGAAAGCATTAAG 540  
QY 541 CAGAACTTCTGGTGTGAGACTCTTTGATGTAAGTGAAGAGGCTTCAAGAGCAAGAG 600  
DB 541 CAGAACTTCTGGTGTGAGACTCTTTGATGTAAGTGAAGAGGCTTCAAGAGCAAGAG 600  
QY 601 GGAAGATGTTGTTTAAACCAACCTTACATCACTACTCAAGAGCGTGCATACCGAA 660  
DB 601 GGAAGATGTTGTTTAAACCAACCTTACATCACTACTCAAGAGCGTGCATACCGAA 660  
QY 661 CGAGAGGGGGGTGAGAGCTGCCAGGTGGGGCTTTGGCATCTCATTCGATCCGTGG 720  
DB 661 CGAGAGGGGGGTGAGAGCTGCCAGGTGGGGCTTTGGCATCTCATTCGATCCGTGG 720  
QY 721 CTTCTTCTTCATCTACAGCTCTCAACAGATATTCAGAAATCCAGATGGGCTGCCA 780  
DB 721 CTTCTTCTTCATCTACAGCTCTCAACAGATATTCAGAAATCCAGATGGGCTGCCA 780  
QY 781 AAATTCCAACAGCTGATTAAGGTGGAAGATGCAAGAAATGATGCAAGAAATGAGCTTC 840  
DB 781 AAATTCCAACAGCTGATTAAGGTGGAAGATGCAAGAAATGATGCAAGAAATGAGCTTC 840  
QY 841 ATGGGATCAAAATTTGATTCATTCAGCTAAAGATGGGGGAAAGACCTACAGATACGATT 900  
DB 841 ATGGGATCAAAATTTGATTCATTCAGCTAAAGATGGGGGAAAGACCTACAGATACGATT 900  
QY 901 CTTTCAACCTGTAGCAAGATCACTGGGAGCAAAATTCAGAAACAGGTTGTACTGTCA 960  
DB 901 CTTTCAACCTGTAGCAAGATCACTGGGAGCAAAATTCAGAAACAGGTTGTACTGTCA 960  
QY 961 GTGACATCTGAGAGCTGGATGTTGGGAGGGGTGCATGATGATGGGGGTGGAGCT 1020  
DB 961 GTGACATCTGAGAGCTGGATGTTGGGAGGGGTGCATGATGATGGGGGTGGAGCT 1020  
QY 1021 TTATATCATGGGAGAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGATC 1080  
DB 1021 TTATATCATGGGAGAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGATC 1080  
QY 1081 TGGGCTGTGCTCTGTGAGCTGCAAGAAAGACAGAGTGGATGGTCTTCAATATATC 1140  
DB 1081 TGGGCTGTGCTCTGTGAGCTGCAAGAAAGACAGAGTGGATGGTCTTCAATATATC 1140  
QY 1141 AGTTACCAAGATTAATTTTCAACTGACGTGGGATGAGGTGACGAGAACT 1200  
DB 1141 AGTTACCAAGATTAATTTTCAACTGACGTGGGATGAGGTGACGAGAACT 1200  
QY 1201 TCTTACCACTGGGCTGCAATTCAGTGGCAGTAAAGGCGCAGGCGCATATGAGGAGG 1260  
DB 1201 TCTTACCACTGGGCTGCAATTCAGTGGCAGTAAAGGCGCAGGCGCATATGAGGAGG 1260  
QY 1261 TTTATGAGCTGTGAGCCCTTCAATATCACTAGTCTCTGAGCCATGAGAGGAGCAG 1320  
DB 1261 TTTATGAGCTGTGAGCCCTTCAATATCACTAGTCTCTGAGCCATGAGAGGAGCAG 1320  
QY 1321 ACATCAACTTTTGAATCAAGCTGAGATGCTGAGGCAAGCTTATGATGACTTATACA 1380  
DB 1321 ACATCAACTTTTGAATCAAGCTGAGATGCTGAGGCAAGCTTATGATGACTTATACA 1380  
QY 1381 AGTATTTCTTCTTCATCACTCCAGAGAGACACATGATGATGATGATGATGATGATG 1440  
DB 1381 AGTATTTCTTCTTCATCACTCCAGAGAGACACATGATGATGATGATGATGATGATG 1440  
QY 1441 TGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
DB 1441 TGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
QY 1501 TGTGCTAGGTTCTTAAAGAGTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1560  
DB 1501 TGTGCTAGGTTCTTAAAGAGTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1560  
QY 1561 CTGGGCTGCAACTTTGAAAACCTCTTTCACATTAACAAATTCATCAATCTTCA 1620



Db 1561 CTGGGCTGCAACTTTGGAAAACTCCTCTGCATACATTAATTCATTCATCTCA 1620  
 Qy 1621 AAGCAACCTTATTTTCATGCTTTCTGTATATTTCTTGATCTTCCAAATTCCT 1680  
 Db 1621 AAGCAACCTTATTTTCATGCTTTCTGTATATTTCTTGATCTTCCAAATTCCT 1680  
 Qy 1681 GATCTTAGAAAAAGAAATCATTTCTCCCTCCCTCCACACATGATCAATATGTA 1740  
 Db 1681 GATCTTAGAAAAAGAAATCATTTCTCCCTCCCTCCACACATGATCAATATGTA 1740  
 Qy 1741 GGGATTACAGTGGGGGCAATTTCTTTATATCACTCTTAAAAAATGTTTCCATTAA 1800  
 Db 1741 GGGATTACAGTGGGGGCAATTTCTTTATATCACTCTTAAAAAATGTTTCCATTAA 1800  
 Qy 1801 AGTAAACCTTATTAATTTTGGAAAGATCTCTGAAAAAATTTTAAAAA 1851  
 Db 1801 AGTAAACCTTATTAATTTTGGAAAGATCTCTGAAAAAATTTTAAAAA 1851  
 RESULT 7  
 ID AB271453 standard; cDNA; 1863 BP.  
 AC AB271453;  
 XX 04-APR-2003 (first entry)  
 DT  
 XX Secreted protein-encoding gene 142 cDNA clone HRAC135, SEQ ID NO:274.  
 DE  
 XX Human; secreted protein; digestive disorder; gastrointestinal disorder;  
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;  
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;  
 KW immune disorder; inflammation; infection; wound healing; drug screening;  
 KW chromosome identification; chromosome mapping; cytostatic; gene therapy;  
 KW antiinflammatory; immunosuppressive; vulnery; chromosome 8q22.2;  
 KW gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200276488-A1.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 19-MAR-2002; 2002MO-US08276.  
 XX  
 PR 21-MAR-2001; 2001US-277340P.  
 PR 19-JUL-2001; 2001US-306171P.  
 PR 13-NOV-2001; 2001US-331287P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 P1 Rosen CA, Ruben SM,  
 XX  
 DR WPI; 2003-029900/02.  
 DR P-PSDB; ABR00274.  
 XX  
 FT New human secreted proteins and nucleic acids, useful for detecting,  
 FT preventing, diagnosing, prognosticating, treating and/or ameliorating  
 FT e.g. gastrointestinal diseases and disorders, or cancers -  
 XX  
 XX Claim 21; Page 918; 1216p; English.  
 XX  
 CC AB271190-AB271478 represent cDNAs corresponding to 178 human secreted  
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.  
 CC AB271479-AB271540 represent human secreted protein genomic fragments. The  
 CC invention also encompasses antibodies specific for the secreted proteins,  
 CC the use of the secreted proteins in drug screening, and recombinant  
 CC vectors and host cells comprising a nucleic acid of the invention. The  
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
 CC fragments are useful for the secreted proteins, and modulators of protein  
 CC activity are useful for diagnosing, treating, ameliorating or preventing  
 CC digestive disorders. Such conditions include disorders of the mouth,  
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary

CC tract and pancreas, and include cancers of these organs and tissues. The  
 CC secreted proteins and their nucleic acids may also be used in the  
 CC treatment of immune disorders, inflammation, infection,  
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids  
 CC of the invention may be used for chromosome identification, chromosome  
 CC mapping, in gene therapy, for identifying individuals from minute  
 CC biological samples, as hybridisation probes, and as molecular weight  
 CC markers. The present sequence represents a human secreted protein-  
 CC encoding cDNA clone of the invention.  
 XX  
 SQ Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;  
 Query Match 99.7%; Score 1846.2; DB 25; Length 1863;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1848; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 GGCTAGCCCGAGCCTTAGTCTGAGAGCCGCTCCGCTCCGCTCAGAGCCGCCCTTA 60  
 Db 1 GACTAGCCCGGAGCCTTAGTCTGAGAGCCGCTCCGCTCCGCTCAGAGCCGCCCTTA 60  
 Qy 61 TCAGATTATCTTAACAAGAAAAACCACTGGAAGAAAAAATGAATTCCTATCTTCGCAT 120  
 Db 61 TCAGATTATCTTAACAAGAAAAACCACTGGAAGAAAAAATGAATTCCTATCTTCGCAT 120  
 Qy 121 TTTTCGGTGGTGTCACTTTTATCCCTGCTCTGGGAAAAGCTATATGCAAGATGCA 180  
 Db 121 TTTTCGGTGGTGTCACTTTTATCCCTGCTCTGGGAAAAGCTATATGCAAGATGCA 180  
 Qy 181 TCTCTAAGAGACTTTTGAAGAAATTAAGAAATATGCCAGCTGTGAGATGTGCTA 240  
 Db 181 TCTCTAAGAGACTTTTGAAGAAATTAAGAAATATGCCAGCTGTGAGATGTGCTA 240  
 Qy 241 AAGAAATCATCAACTGCTGTTATGTTAAAGCCAAACAGATCCATGAGGATTTG 300  
 Db 241 AAGAAATCATCAACTGCTGTTATGTTAAAGCCAAACAGATCCATGAGGATTTG 300  
 Qy 301 CACTTCTGTTGATATCTTGGACCCAGACTGAGTGGCTCCAGAACCTTAGAAAAAGCA 360  
 Db 301 CACTTCTGTTGATATCTTGGACCCAGACTGAGTGGCTCCAGAACCTTAGAAAAAGCA 360  
 Qy 361 TCCAATTATGTATACCAAAACCTGAGCAAGATGGCTGAGAAAGTTCACTGAGCCAG 420  
 Db 361 TCCAATTATGTATACCAAAACCTGAGCAAGATGGCTGAGAAAGTTCACTGAGCCAG 420  
 Qy 421 TGAGAAATACCCCACTGGGAGAGGAGAAAGAAATGAGTGTGAGTGGAGCAAGAATTC 480  
 Db 421 TGAGAAATACCCCACTGGGAGAGGAGAAAGAAATGAGTGTGAGTGGAGCAAGAATTC 480  
 Qy 481 ATTAAGATAGCCATCTGGGATCTTGGACAGACATTTGGGACTCTCCAGAAAGCAATTAC 540  
 Db 481 ATTAAGATAGCCATCTGGGATCTTGGACAGACATTTGGGACTCTCCAGAAAGCAATTAC 540  
 Qy 541 CAGAAGTTCTGGTGGTGAACCTTTTCATGAACTGCAAGAGAGGCTCAGAGCAAGAG 600  
 Db 541 CAGAAGTTCTGGTGGTGAACCTTTTCATGAACTGCAAGAGAGGCTCAGAGCAAGAG 600  
 Qy 601 GGAAGATTTGTTTATTAACAACCTTACATCACTCAAGAGAGGCTGAATACGGA 660  
 Db 601 GGAAGATTTGTTTATTAACAACCTTACATCACTCAAGAGAGGCTGAATACGGA 660  
 Qy 661 CGCAGAGGCGGTGAAGAGCTCCAAAGTGGGGCTTTGGACTCTCAATTCGATCCGTGG 720  
 Db 661 CGCAGAGGCGGTGAAGAGCTCCAAAGTGGGGCTTTGGACTCTCAATTCGATCCGTGG 720  
 Qy 721 CTTCTTCTCATCTACAGTCTTCAACAGAGTATTCAGAAATACAGAGATGGCTGCCA 780  
 Db 721 CTTCTTCTCATCTACAGTCTTCAACAGAGTATTCAGAAATACAGAGATGGCTGCCA 780  
 Qy 781 AAATTCAGACGCTGTATTAACGTTGAGAAAGTGAAGAAATGATCAAGAAATGCTTCTC 840  
 Db 781 AAATTCAGACGCTGTATTAACGTTGAGAAAGTGAAGAAATGATCAAGAAATGCTTCTC 840  
 Qy 841 ATGGATCAAAATGTATCATCAGCTAAAGATGGGGGCAAGACCTACCACTACTGAT 900

Db	841	ATGGGATCAAATTGTCAATTCAAGCTTAAGATGGGGGCAAAAGACTTACCAAGTACTGATT	900
Qy	901	CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGTTGTACTGTCA	960
Db	901	CCTTCAACACTGTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGTTGTACTGTCA	960
Qy	961	GTGACATCTGGACAGCTGGGATGTTGGGCAAGGGTGCATGATGATAGGGGTGGAGCCT	1022
Db	961	GTGACATCTTGGACAGCTGGGATGTTGGGCAAGGGTGCATGATGATAGGGGTGGAGCCT	1022
Qy	1021	TTATATCATGGGAAGCACTCTCACTTATTTAAAGATCTTGGGCTGGCTCAAGAGACTC	1080
Db	1021	TTATATCATGGGAAGCACTCTCACTTATTTAAAGATCTTGGGCTGGCTCAAGAGACTC	1080
Qy	1081	TGCGGCTGTGTCTCTGGACTGCAAGAAACAAAGGTGGATGTGTGCTTCCAGTATTATC	1144
Db	1081	TGCGGCTGTGTCTCTGGACTGCAAGAAACAAAGGTGGATGTGTGCTTCCAGTATTATC	1144
Qy	1141	AGTTACACAAGTAAATATTTTCCACTACGTCTGTATGAGAGCTGACGCAAGAACCT	1200
Db	1141	AGTTACACAAGTAAATATTTTCCACTACAGTCTGTATGAGAGCTTACGCAAGAACCT	1200
Qy	1201	TCTTACCACCTGGGCTGCATTTCACTGCAAGTAAAGGCCACATCATGAGAGAG	1266
Db	1201	TCTTACCACCTGGGCTGCATTTCACTGCAAGTAAAGGCCACATCATGAGAGAG	1266
Qy	1261	TTATGAGCTGTGTGAGCCCTCAATATCACTCAGTCTGTAGGCATGAGAAAGGACAG	1322
Db	1261	TTATGAGCTGTGTGAGCCCTCAATATCACTCAGTCTGTAGGCATGAGAAAGGACAG	1322
Qy	1321	ACATCAACTTTTGTGATCCCAAGCTGAGTGTCTGAGCCAGTCTACTTATGATGATTAACA	1388
Db	1321	ACATCAACTTTTGTGATCCCAAGCTGAGTGTCTGAGCCAGTCTACTTATGATGATTAACA	1388
Qy	1381	AGTATTTCTTCTTCATCATCACTCCACGGAGACACATGATCTGTGATCCCAAGCAGA	1444
Db	1381	AGTATTTCTTCTTCATCATCACTCCACGGAGACACATGATCTGTGATCCCAAGCAGA	1444
Qy	1441	TGAATGTTGTCTGTCTGTGTGGGCTGTGTTCTTATGTTGTTGCAGACATGGAAGAA	1500
Db	1441	TGAATGTTGTCTGTCTGTGTGGGCTGTGTTCTTATGTTGTTGCAGACATGGAAGAA	1500
Qy	1501	TGCTGTCTGTAGTCTTGAAGAACATGAAGAACGTTTTCATGCTTCTGTGCGCAGGAATC	1566
Db	1501	TGCTGTCTGTAGTCTTGAAGAACATGAAGAACGTTTTCATGCTTCTGTGCGCAGGAATC	1566
Qy	1561	CTGGGCTGCAACTTTGGAAAACTCTCTGCACAAACAATTTCATCAATCTTCA	1622
Db	1561	CTGGGCTGCAACTTTGGAAAACTCTCTGCACAAACAATTTCATCAATCTTCA	1622
Qy	1621	AAGCAACACTGTATTCAATGCTTCTGTATTATATCTTTCTTGATACCTTCCAAATCTCT	1680
Db	1621	AAGCAACACTGTATTCAATGCTTCTGTATTATATCTTTCTTGATACCTTCCAAATCTCT	1680
Qy	1681	GATTCCTGAAGAAAGGAATCATTTCCCTCCCTCCACACATGAATCAATATGGTA	1744
Db	1681	GATTCCTGAAGAAAGGAATCATTTCCCTCCCTCCACACATGAATCAATATGGTA	1744
Qy	1741	GGGATTCACGTGGGGGCAATTTCTTATATACACTCTTAAAAAACATGTGTTCCACTTAA	1800
Db	1741	GGGATTCACGTGGGGGCAATTTCTTATATACACTCTTAAAAAACATGTGTTCCACTTAA	1800
Qy	1801	AGTAAACCTTATTAATTTTTTGGAGATCTTGAAAAAATTTTTTTTTTTTTTTT	1851
Db	1801	AGTAAACCTTATTAATTTTTTGGAGATCTTGAAAAAATTTTTTTTTTTTTTTT	1851

RESULT	8
AAZ98139	
ID	AAZ98139 standard; cDNA; 1923 BP.
XX	
AC	AAZ98139;

XX 11-MAY-2000 (first entry)  
DT  
XX  
DE Human signal peptide containing protein HSPB-31 cDNA SEQ ID NO:165.  
XX  
KW Human; signal peptide-containing protein; HSPB; diagnosis; cancer;  
inflammation; cardiovascular disease; anticancer; anti-inflammatory;  
antimicrobial; neurotropic; neuroprotective; cardiovascular; hepatotropic;  
antiaesthetic; gene therapy; cell proliferation; neurological disorder;  
reproductive disorder; developmental disorder; arteriosclerosis;  
cirrhosis; porriasis; acquired immune deficiency syndrome; anaemia;  
asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;  
Parkinson's disease; Huntington's disease; ovulatory defect;  
muscular dystrophy; ss.  
XX  
OS Homo sapiens.  
XX  
PM WO200000610-A2.  
XX  
PD 06-JAN-2000.  
XX  
PF 25-JUN-1999; 99WO-US14484.  
XX  
PR 26-JUN-1998; 98US-0090762.  
PR 31-JUL-1998; 98US-0094983.  
PR 01-OCT-1998; 98US-0102686.  
PR 11-DEC-1998; 98US-0112129.  
XX  
XX (INCY-) INCYTE PHARM INC.  
PA  
P1 Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;  
P1 Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JT;  
P1 Bandman O;  
XX  
DR WPI: 2000-160673/14.  
XX  
P-PSDB: AAY87254.  
XX  
DR New human signal peptide-containing proteins useful in treatment,  
PT prevention and diagnosis of e.g. cancer, inflammation and  
PT cardiovascular disease  
XX  
PS Claim 9; Page 269-270; 327pp; English.  
XX  
XX AA298109 to AA298242 encode AAY87224 to AAY87357 which represent the  
CC human signal peptide-containing proteins HSPB-1 to HSPB-134. HSPBs have  
CC anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatotropic,  
CC neuroprotective, cardiovascular and antiaesthetic activities, and can  
CC be used in gene therapy. HSPBs can be used to treat or prevent disorders  
CC associated with decreased activity or function of HSPB. Antagonists of  
CC HSPB are used to treat or prevent disorders associated with increased  
CC activity or function of HSPB. Such diseases include cell proliferation  
CC (including cancer), inflammation, cardiovascular, neurological,  
CC reproductive or developmental disorders, (e.g. arteriosclerosis,  
CC cirrhosis, porriasis, acquired immune deficiency syndrome, anaemia,  
CC asthma, Crohn's disease, microbial or other infections, congestive or  
CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's  
CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSPB  
CC nucleic acids can be used for the recombinant production of HSPB, for  
CC detecting HSPB in standard hybridisation and amplification assays (for  
CC diagnosis and monitoring), in gene therapy, as antisense,  
CC triplex-forming or ribozyme therapeutics, for detecting related sequences  
CC or genetic variations, and for chromosomal mapping. HSPB are also used to  
CC raise specific antibodies (Ab) and to screen for agonists and  
CC antagonists (potential therapeutic agents). Ab are used to diagnose, or  
CC monitor, HSPB-related diseases (in usual immunoassays), as therapeutic  
CC antagonists, in competitive drug screens, and for purification of HSPB  
CC from natural sources.  
XX  
SQ Sequence 1923 BP; 538 A; 439 C; 471 G; 475 T; 0 other;  
Query Match 99.0%; Score 1833.4; DB 21; Length 1923;  
Best Local Similarity 99.7%; Pred. NO. 0;  
Matches 1837; Conservative 6; Indels 0; Gaps 0;

Oy 1 GGCTAGGCGGAGCTTAGTCTCTGGAGCCGCTCCGTGCGCCGCTGAGAGCCGCCCTA 60  
 Db 81 GGCTAGGCGGAGCTTAGTCTCTGGAGCCGCTCCGTGCGCCGCTGAGAGCCGCCCTA 140  
 Oy 61 TCAGATTATCTTAAACAAGAAAACAAGTGAATAAATAATTCCTTATCTTCCGAT 120  
 Db 141 TCAGATTATCTTAAACAAGAAAACAAGTGAATAAATAATTCCTTATCTTCCGAT 200  
 Oy 121 TTTTGGGTGGTGTCACTTTATATCCCTGTCTCTGGGAAAGCTATATGCAAGATGGCA 180  
 Db 201 TTTTGGGTGGTGTCACTTTATATCCCTGTCTCTGGGAAAGCTATATGCAAGATGGCA 260  
 Oy 181 TCTTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGGAGATGTTGCTA 240  
 Db 261 TCTTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGGAGATGTTGCTA 320  
 Oy 241 AAGCAATCATCAACCTAGCTGTATATGTTAAAGCCAGAAAGATCTATAGAGCGATTGG 300  
 Db 321 AAGCAATCATCAACCTAGCTGTATATGTTAAAGCCAGAAAGATCTATAGAGCGATTGG 380  
 Oy 301 CACTTCTGGTGTATCTGTGGAGCCCAAGTGTGCTCCAAAGAACTTGAAGAAAGCCA 360  
 Db 381 CACTTCTGGTGTATCTGTGGAGCCCAAGTGTGCTCCAAAGAACTTGAAGAAAGCCA 440  
 Oy 361 TCCAAATATATGTAACCAAAACCTGCAAGCAAGATGGGCTGGAGAAAGTTGACCTGGAGCGAG 420  
 Db 441 TCCAAATATATGTAACCAAAACCTGCAAGCAAGATGGGCTGGAGAAAGTTGACCTGGAGCGAG 500  
 Oy 421 TGAGAAATACCCCACTGGAGAGAGGAGAGAAATCAAGCTGTGTATGCTGAGGCCAAGAAATTC 480  
 Db 501 TGAGAAATACCCCACTGGAGAGAGGAGAGAAATCAAGCTGTGTATGCTGAGGCCAAGAAATTC 560  
 Oy 481 ATAAATATAGCCATCTGGGTCTTTGGAGAGCAATGGGACTCTCCAAAGGCAATTAACAG 540  
 Db 561 ATAAATATAGCCATCTGGGTCTTTGGAGAGCAATGGGACTCTCCAAAGGCAATTAACAG 620  
 Oy 541 CAGAGATTCTGGTGTGACCTCTTCCATGAACTGCAAGAGAGGAGGCTCAGAAAGCAAGAG 600  
 Db 621 CAGAGATTCTGGTGTGACCTCTTCCATGAACTGCAAGAGAGGAGGCTCAGAAAGCAAGAG 680  
 Oy 601 GGAAGATTGTGTTTATTAACAACCTTACATCACTCAAGAGAGCGGTGCAATACCGAA 660  
 Db 681 GGAAGATTGTGTTTATTAACAACCTTACATCACTCAAGAGAGCGGTGCAATACCGAA 740  
 Oy 661 CGCAGGGGCGGTGGAAGCTGCCAAGGTGGGCTTTGGCATCTCATTCGATCCGTGG 720  
 Db 741 CGCAGGGGCGGTGGAAGCTGCCAAGGTGGGCTTTGGCATCTCATTCGATCCGTGG 800  
 Oy 721 CCTCCTTCCATCTACAGTCTCTCAACAGGTAATCAAGGAATACGAGATGGCGGCCA 780  
 Db 801 CCTCCTTCCATCTACAGTCTCTCAACAGGTAATCAAGGAATACGAGATGGCGGCCA 860  
 Oy 781 AAATTCACAAGCCCTGTATTAACGTGGAAGATGCAAGAAATGATCAAGAAATGCGTCTC 840  
 Db 861 AAATTCACAAGCCCTGTATTAACGTGGAAGATGCAAGAAATGATCAAGAAATGCGTCTC 920  
 Oy 841 ATGGATCAAAATTTGTCAATTCAGCTAAAGATGGGGCAAAAGACTTACCCAGATCTGATT 900  
 Db 921 ATGGATCAAAATTTGTCAATTCAGCTAAAGATGGGGCAAAAGACTTACCCAGATCTGATT 980  
 Oy 901 CCTTCAACACTGTACAGAGATCACTGGGAGCAAAATTCACAAAGATGTAAGTCACTGCTCA 960  
 Db 981 CCTTCAACACTGTACAGAGATCACTGGGAGCAAAATTCACAAAGATGTAAGTCACTGCTCA 1040  
 Oy 961 GTGACATCTGGAGAGCTGGAGATGTTGGGCAAGGCTCCATGATGATGGCGGTGGAGCCT 1020  
 Db 1041 GTGACATCTGGAGAGCTGGAGATGTTGGGCAAGGCTCCATGATGATGGCGGTGGAGCCT 1100  
 Oy 1021 TTATATCATGGAGAGCACTCTCATTTATTAAGATCTTGGGCTGCGTCCAAAGAGCACTC 1080  
 Db 1101 TTATATCATGGAGAGCACTCTCATTTATTAAGATCTTGGGCTGCGTCCAAAGAGCACTC 1160

Oy 1081 TGCGGCTGTGCTCTGTGACCTGACAGAGAACAAGGTGAGATTGTGCTTCCAGATTATATC 1140  
 Db 1161 TGCGGCTGTGCTCTGTGACCTGACAGAGAACAAGGTGAGATTGTGCTTCCAGATTATATC 1220  
 Oy 1141 AGTTACACAGGTAATATTTTCCAACTACAGTCTGTGTATGAGATCTGACCGAGAACTT 1200  
 Db 1221 AGTTACACAGGTAATATTTTCCAACTACAGTCTGTGTATGAGATCTGACCGAGAACTT 1280  
 Oy 1201 TCTTACCCACTGGGCTGCAATTCAGTGGCAGTGAAGAGGCGGCAATCAATGAGGAGG 1260  
 Db 1281 TCTTACCCACTGGGCTGCAATTCAGTGGCAGTGAAGAGGCGGCAATCAATGAGGAGG 1340  
 Oy 1261 TTATAGAGCTGTGCTGAGCCCTCAATATCACTCAGGTCTTGAAGCATGAGAGAGGACAG 1320  
 Db 1341 TTATAGAGCTGTGCTGAGCCCTCAATATCACTCAGGTCTTGAAGCATGAGAGAGGACAG 1400  
 Oy 1321 ACATCAACTTTTGGATCCAGCTGAGAGTCTGAGAGCCAGTCTACTGATGATCTTATACA 1380  
 Db 1401 ACATCAACTTTTGGATCCAGCTGAGAGTCTGAGAGCCAGTCTACTGATGATCTTATACA 1460  
 Oy 1381 AGTATTTCTTCTTGCATCACTCCAGAGAGACACATGACATGATGATGATCAAGAGAGA 1440  
 Db 1461 AGTATTTCTTCTTGCATCACTCCAGAGAGACACATGACATGATGATGATCAAGAGAGA 1500  
 Oy 1441 TGAATGTTGCTGCTGTGCTGTGAGGCTGTGTTCTTATATGTTGTCAGACATGAGAGAA 1500  
 Db 1521 TGAATGTTGCTGCTGTGCTGTGAGGCTGTGTTCTTATATGTTGTCAGACATGAGAGAA 1580  
 Oy 1501 TGCTGCTAGGTCCTTGAAGACATGTAAGAAAGAAAGTTTATGCTTCTGCGCAGGAATC 1560  
 Db 1581 TGCTGCTAGGTCCTTGAAGACATGTAAGAAAGAAAGTTTATGCTTCTGCGCAGGAATC 1640  
 Oy 1561 CTGGGTCTGCAACTTTGAAAGAACTCTCTTCAATCAATCAATCAATCAATCAATCAATCA 1620  
 Db 1641 CTGGGTCTGCAACTTTGAAAGAACTCTCTTCAATCAATCAATCAATCAATCAATCAATCA 1700  
 Oy 1621 AAGCAAACTTATTTATCAATGCTTCTGTATATCTTCTGATATCTTCCAAATTTCTCT 1680  
 Db 1701 AAGCAAACTTATTTATCAATGCTTCTGTATATCTTCTGATATCTTCCAAATTTCTCT 1760  
 Oy 1681 GATTCTGAGAAAAGAAATCATTTCTCCCTCCCTCCCAACATGAAATCAATATGCTA 1740  
 Db 1761 GATTCTGAGAAAAGAAATCATTTCTCCCTCCCTCCCAACATGAAATCAATATGCTA 1820  
 Oy 1741 GGGATTTACAGTGGGCGCACTTTCTTATATCACTCTTAAACAACTGTTTCCACTTAAA 1800  
 Db 1821 GGGATTTACAGTGGGCGCACTTTCTTATATCACTCTTAAACAACTGTTTCCACTTAAA 1880  
 Oy 1801 AGTAAACACTTAATTAATTTTGGAGAGTCTGTAAGAAAAA 1843  
 Db 1881 AGTAAACACTTAATTAATTTTGGAGAGTCAAAAAA 1923

RESULT 9  
 AA258313  
 ID AA258313 standard; cDNA; 1884 BP.  
 XX  
 AA258313;  
 DT 08-MAY-2000 (first entry)  
 XX  
 DE Human peptidase NAALAD-ase IV cDNA.  
 XX  
 KW NAALAD-ase IV; N-acetylated alpha-linked acidic dipeptidase; human;  
 KW chromosome 8q21.3; prostate cancer; neurodegenerative disease;  
 KW Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;  
 KW peripheral neuropathy; Huntington's disease; acute brain injury;  
 KW multiple sclerosis; peripheral nerve trauma; ischemia; dementia;  
 KW gene therapy; diagnosis; noctropic; neuroprotective; neuroleptic;  
 KW antiparkinsonian; anticonvulsant; vasotropic; ss.  
 OS Homo sapiens.  
 XX

FH Key Location/Qualifiers  
 FT CDS 149..1567  
 FT /+tag= a  
 XX  
 XX WO200004157-A2.  
 XX  
 XX 27-JAN-2000.  
 XX  
 XX 14-JUL-1999; 99WO-GB02241.  
 XX PF  
 XX 14-JUL-1998; 98GB-0015284.  
 XX PR  
 XX (JANC) JANSSEN PHARM NV.  
 XX PA  
 XX Pangalos M, Neefs JEFM, Peeters DCG;  
 XX WPI: 2000-182424/16.  
 XX P-PSDB; AAY58879.  
 DR  
 XX  
 XX New human N-acetylated alpha-linked acidic dipeptidases for treating  
 PT neural disorders e.g. Alzheimer's disease, schizophrenia and  
 PT Parkinson's disease  
 XX  
 XX Claim 7; Fig 5; 95pp; English.  
 PS  
 XX  
 XX The present sequence is that of cDNA coding for human  
 CC N-acetylated alpha-linked acidic dipeptidase IV (NALAD-ase IV,  
 CC see AAY58879). The cDNA was obtained from a gall bladder cDNA  
 CC library. Analysis of the open reading frame predicts a type II  
 CC integral membrane protein with 5 potential N-glycosylation sites.  
 CC The NALAD-ase II gene was mapped to chromosome 8q21.3.  
 CC NALAD-ase IV expression was low in all tissues examined by RT-PCR.  
 CC The invention provides human NALAD-ase I, II and IV polypeptides,  
 CC cDNAs, antisense nucleic acids, vectors, host cells, transgenic  
 CC organisms, antagonists and agonists. These are useful for treating  
 CC neural disorders such as Alzheimer's disease, schizophrenia, ALS,  
 CC Parkinson's disease, peripheral neuropathy, Huntington's disease,  
 CC acute brain injury, multiple sclerosis, exposure to neurotoxins,  
 CC peripheral nerve trauma, ischemia or dementia (claimed). Nucleic  
 CC acids can also be used for gene therapy and for genetic screening  
 CC of predisposition to disorders associated with NALAD-ase.  
 XX  
 XX Sequence 1884 BP; 520 A; 432 C; 458 G; 474 T; 0 other;  
 SQ  
 Query Match 98.9%; Score 1830.8; DB 21; Length 1884;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1832; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db  
 411 TCCTAAATTAATGATACCAAAACCTGAGCAAGATGGCTGAGAAAGTTCACTGAGCCAG 470  
 Qy  
 421 TGAATAATACCCCTGGGAGAGGGAGAAATAGCTGTGATCTGTGAGCAAGATTC 480  
 Db  
 471 TGAATAATACCCCTGGGAGAGGGAGAAATAGCTGTGATCTGTGAGCAAGATTC 530  
 Qy  
 481 ATAAGATAGCCATCTGGGTCTTGGAGCAGCATTTGGGACTCTCCAGAAAGCATTCAG 540  
 Db  
 531 ATAAGATAGCCATCTGGGTCTTGGAGCAGCATTTGGGACTCTCTCCAGAAAGCATTCAG 590  
 Qy  
 541 CAGAGTTCTGTGTGATGACCTCTTTGATGAACTGACAGAGAGGGCTCAGAGCAAGAG 600  
 Db  
 591 CAGAGTTCTGTGTGATGACCTCTTTGATGAACTGACAGAGAGGGCTCAGAGCAAGAG 650  
 Qy  
 601 GGAAGATTTGTTGTTTAAACCAACTTACATCACTCTCAGAGACGGTGCATAACCGAA 660  
 Db  
 651 GGAAGATTTGTTGTTTAAACCAACTTACATCACTCTCAGAGACGGTGCATAACCGAA 710  
 Qy  
 661 CGCAGGGGGCGGTGAGAGCTGCCAAGGTGGGGCTTGGCATCTCTGATTCGATCCGTGG 720  
 Db  
 711 CGCAGGGGGCGGTGAGAGCTGCCAAGGTGGGGCTTGGCATCTCTGATTCGATCCGTGG 770  
 Qy  
 721 CCTCTTCTCATCTTACAGTCTTACACAGGTATTCAGAAATACAGATGCGCTGCCA 780  
 Db  
 771 CCTCTTCTCATCTTACAGTCTTACACAGGTATTCAGAAATACAGATGCGCTGCCA 830  
 Qy  
 781 AATATTCACAGCTGTATTAAGGTGAGAAATGATGCAAGAAATGATGCAAGATGCTTCTC 840  
 Db  
 831 AGATTTCAACAGCTGTATTAAGGTGAGAAATGATGCAAGAAATGATGCAAGATGCTTCTC 890  
 Qy  
 841 ATGGGATCAAAATTTGATTCAGTAAAGATGGGGGCAAGAGCTTACCAGATCTGATT 900  
 Db  
 891 ATGGGATCAAAATTTGATTCAGTAAAGATGGGGGCAAGAGCTTACCAGATCTGATT 950  
 Qy  
 901 CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGTTGTACTGTGTCA 960  
 Db  
 951 CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGTTGTACTGTGTCA 1010  
 Qy  
 961 GTGACATCTGAGACAGCTGGGATTTGGGAGGGTGCATGATGATGGCGGTGGAGCTT 1020  
 Db  
 1011 GTGACATCTGAGACAGCTGGGATTTGGGAGGGTGCATGATGATGGCGGTGGAGCTT 1070  
 Qy  
 1021 TTATATCATGGGAGCACTCTCACTTAATAAGATCTTGGGCTGGCTCAAGAGGACTC 1080  
 Db  
 1071 TTATATCATGGGAGCACTCTCACTTAATAAGATCTTGGGCTGGCTCAAGAGGACTC 1130  
 Qy  
 1081 TGCAGCTGTGCTCTGTGACTGACAGAAAGCAAGGTGGAGTTGTGCTTCCAGATTAATC 1140  
 Db  
 1131 TGCAGCTGTGCTCTGTGACTGACAGAAAGCAAGGTGGAGTTGTGCTTCCAGATTAATC 1190  
 Qy  
 1141 AGTTACCAAGGTAAATATTTCCACTACAGTCTGTGATGAGATCTGAGCGCAGAACTT 1200  
 Db  
 1191 AGTTACCAAGGTAAATATTTCCACTACAGTCTGTGATGAGATCTGAGCGCAGAACTT 1250  
 Qy  
 1201 TCTTACCACTGGGCTGCAATTCACGTGCAGTGAAGGCGAGGCGCATATGAGAGAG 1260  
 Db  
 1251 TCTTACCACTGGGCTGCAATTCACGTGCAGTGAAGGCGAGGCGCATATGAGAGAG 1310  
 Qy  
 1261 TTATGAGCTGTGCTGACAGCCCTCAATATCACTCAGTCTGTGAGCATGAGAGAGGACAG 1320  
 Db  
 1311 TTATGAGCTGTGCTGACAGCCCTCAATATCACTCAGTCTGTGAGCATGAGAGAGGACAG 1370  
 Qy  
 1321 ACATCAACTTTTGGATCCAAAGCTGAGTGGCTGAGCGCAGTCTTATGATGACTTATACA 1380  
 Db  
 1371 ACATCAACTTTTGGATCCAAAGCTGAGTGGCTGAGCGCAGTCTTATGATGACTTATACA 1430  
 Qy  
 1381 AGTATTTCTTCTTCATCACTCCACGAGAGACATGATGATGATGATGATGATGATGATGAT 1440  
 Db  
 1431 AGTATTTCTTCTTCATCACTCCACGAGAGACATGATGATGATGATGATGATGATGATGAT 1490  
 Qy  
 1441 TGAATGTGCTGCTGCTGTTGGGCTGTTTCTTATGTTGTTGACAGACATGAGAGAA 1500



```

Db      754 CCTCTCTCCATCTACAGTCTCTCACAGGATTTACAGAAATACAGAGATGGCTGCCA 813
Qy      761 AAATTCACACGCGCTGATTTACGGTGAAGATGCAAGAAATGTCAGAAATGGCTTCTC 840
Db      814 AGATTCACACGCGCTGATTTACGGTGAAGATGCAAGAAATGTCAGAAATGGCTTCTC 873
Qy      841 ATGGATCAAAATTTGTCATTCAGCTTAAAGATGGGGGCAAGACCTACCCAGATCTGATT 900
Db      874 ATGGGATCAAAATTTGTCATTCAGCTTAAAGATGGGGGCAAGACCTACCCAGATCTGATT 933
Qy      901 CTTTCAACTGTGACAGAGATCACTGGAGCAAAATATCCAGAAACAGTTTACTGTCTCA 960
Db      934 CCTTCAACACTGTGACAGAGATCACTGGAGCAAAATATCCAGAAACAGTTTACTGTCTCA 993
Qy      961 GTGACATCTGGACAGCTGGGATGTTGGGCGAGGCTGCATGATGATGCGGTGGAGCT 1020
Db      994 GTGACATCTGGACAGCTGGGATGTTGGGCGAGGCTGCATGATGATGCGGTGGAGCT 1053
Qy      1021 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTC 1080
Db      1054 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTC 1113
Qy      1081 TGGCGCTGTGCTCTGGAAGTGCAGAGAAAGAGGTTGAGTGTGCTTCCAGATTTATC 1140
Db      1114 TGGCGCTGTGCTCTGGAAGTGCAGAGAAAGAGGTTGAGTGTGCTTCCAGATTTATC 1173
Qy      1141 AGTTACACAGGTAATTTTCCAACTTCTGTGATGATGATGATGATGATGATGATGATGAT 1200
Db      1174 AGTTACACAGGTAATTTTCCAACTTCTGTGATGATGATGATGATGATGATGATGATGAT 1233
Qy      1201 TCTTACCACTGGGCTGCAATTCCTGGCAGTGAAGAGGCGCCATCATGGAGGAGG 1260
Db      1234 TCTTACCACTGGGCTGCAATTCCTGGCAGTGAAGAGGCGCCATCATGGAGGAGG 1293
Qy      1261 TTATGAGCTGTGCTGAGGCTGCAATTCCTGAGCTGCTGAGCCATGAGAGAGGAGCAG 1320
Db      1294 TTATGAGCTGTGCTGAGGCTGCAATTCCTGAGCTGCTGAGCCATGAGAGAGGAGCAG 1353
Qy      1321 ACATCACTTTTGGATCCAGCTGAGGCTGCGAGGCGAGTCTTATGATGATGATGATGAT 1380
Db      1354 ACATCACTTTTGGATCCAGCTGAGGCTGCGAGGCGAGTCTTATGATGATGATGATGAT 1413
Qy      1381 AGTATTTCTTCTTCATCACTCCACGAGGACACCATGACTGTCATGAGTCCAAAGCAGA 1440
Db      1414 AGTATTTCTTCTTCATCACTCCACGAGGACACCATGACTGTCATGAGTCCAAAGCAGA 1473
Qy      1441 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTTCTTATGTTGTCAGACATGAGAGAA 1500
Db      1474 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTTCTTATGTTGTCAGACATGAGAGAA 1533
Qy      1501 TGTGCTCTAGGCTCTAGAAACAGTGAAGAAAGAGGTTTCATGCTTGGCCAGGAATC 1560
Db      1534 TGTGCTCTAGGCTCTAGAAACAGTGAAGAAAGAGGTTTCATGCTTGGCCAGGAATC 1593
Qy      1561 CTGGGCTGCAACTTTGGAAGAACTCCTTCACATTAATTCATCCAAATTCATCTCA 1620
Db      1594 CTGGGCTGCAACTTTGGAAGAACTCCTTCACATTAATTCATCCAAATTCATCTCA 1653
Qy      1621 AAGCACAACCTATTTCAATGCTTCTGTTATATCTTCTTGATATCTTCCAAATTTCTCT 1680
Db      1654 AAGCACAACCTATTTCAATGCTTCTGTTATATCTTCTTGATATCTTCCAAATTTCTCT 1713
Qy      1681 GATTCTAGAAAAAGAAATCATCTTCCCTCCCTCCCAACATGAAATCAACATATGTTA 1740
Db      1714 GATTCTAGAAAAAGAAATCATCTTCCCTCCCTCCCAACATGAAATCAACATATGTTA 1773
Qy      1741 GGGATTAAGTGGGGGCAATTTCTTATATCACTTAAATCAATGTTTCCACTTTAA 1800
Db      1774 GGGATTAAGTGGGGGCAATTTCTTATATCACTTAAATCAATGTTTCCACTTTAA 1833
Qy      1801 AGTAAACACTTAATTAATTTTGAAGATCTCTGA 1835

```

```

Db      1834 AGTAAACACTTAATTAATTTTGAAGATCTCTGA 1868
RESULT 11
AB271331
ID AB271331 standard; cDNA; 2077 BP.
XX
AC AB271331;
XX
DT 04-APR-2003 (first entry)
XX
DE Secreted protein-encoding gene 142 cDNA clone HRAU35, SEQ ID NO:152.
XX
KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
KW immune disorder; inflammation; infection; wound healing; drug screening;
KW chromosome identification; chromosome mapping; cytostatic; gene therapy;
KW antiinflammatory; immunosuppressive; vulnerary; chromosome 8q22.2;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN W020027648-A1.
XX
PD 03-OCT-2002.
XX
PF 19-MAR-2002; 2002MO-US08276.
PR 21-MAR-2001; 2001US-277340P.
PR 19-JUL-2001; 2001US-306171P.
PR 13-NOV-2001; 2001US-331287P.
XX
PA (HUMA-) HUMAN GENOME SCT INC.
XX
PI Rosen CA, Ruben SM;
XX
WP1: 2003-029900/02.
DR P-PSDB: ABR00152.
XX
XX
PT New human secreted proteins and nucleic acids, useful for detecting,
PT preventing, diagnosing, prognosticating, treating and/or ameliorating
PT e.g. gastrointestinal diseases and disorders, or cancers -
XX
XX
PS Claim 21; Page 841-842; 1216pp; English.
XX
XX
CC AB271190-AB271478 represent cDNAs corresponding to 178 human secreted
CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
CC AB271479-AB271540 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening, and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody
CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing, treating, ameliorating or preventing
CC digestive disorders. Such conditions include disorders of the mouth,
CC oesophagus, stomach, small intestine, large intestine, liver, biliary
CC tract and pancreas, and include cancers of these organs and tissues. The
CC secreted proteins and their nucleic acids may also be used in the
CC treatment of immune disorders, inflammation, infection,
CC hyperproliferative disorders, and to promote wound healing. Nucleic acids
CC of the invention may be used for chromosome identification, chromosome
CC mapping, in gene therapy, for identifying individuals from minute
CC biological samples, as hybridisation probes, and as molecular weight
CC markers. The present sequence represents a human secreted protein-
CC encoding cDNA clone of the invention.
XX
SQ Sequence 2077 BP; 597 A; 436 C; 482 G; 562 T; 0 other;
Query Match 98.9%; Score 1830.2; DB 25; Length 2077;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1832; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

1 GGTAGGCCGCGAGCTTACTGCTGGAGCGGCTCCGTGCGCGCGCTGACAGCCGCTTA 60  
34 GGGTAGGCGCGAGCTTACTGCTGGAGCGGCTCCGTGCGCGCGCTGACAGCCGCTTA 93  
QY 61 TCAAGTTATCTTAACAAGAAAACCACTGGAAAAAATGAAATTCCTTATCTTCGAT 120  
Db 94 TCAAGTTATCTTAACAAGAAAACCACTGGAAAAAATGAAATTCCTTATCTTCGAT 153  
QY 121 TTTTGGTGGTGTGACCTTTATCCCTGTGCTGGGAAAGCTATATGCAAGATGGCA 180  
Db 154 TTTTGGTGGTGTGACCTTTATCCCTGTGCTGGGAAAGCTATATGCAAGATGGCA 213  
QY 181 TCTCTAAGAGACTTTTGAAGAAATPAAAGAAATAGCCAGCTGTGGAGATGTTGCTA 240  
Db 214 TCTCTAAGAGACTTTTGAAGAAATPAAAGAAATAGCCAGCTGTGGAGATGTTGCTA 273  
QY 241 AAGCAATCATCACTAGCTGTTTATGTTAAGCCCAAGAACGATCTTATGCGCATGG 300  
Db 274 AAGCAATCATCACTAGCTGTTTATGTTAAGCCCAAGAACGATCTTATGCGCATGG 333  
QY 301 CACTTCTGGTGTATCTGTTGGACCCAGCTGAGTGGCTCAAGAACCTGAGAAAAAGCA 360  
Db 334 CACTTCTGGTGTATCTGTTGGACCCAGCTGAGTGGCTCAAGAACCTGAGAAAAAGCA 393  
QY 361 TCCAAATTATGTACCAAAAACCTGACAGCAAGATGGCTGGAGAAAGTTCACTGAGCGAG 420  
Db 394 TCCAAATTATGTACCAAAAACCTGACAGCAAGATGGCTGGAGAAAGTTCACTGAGCGAG 453  
QY 421 TGAAGAAATCCCACTGGGAGAGGGAAGAAATCAGCTGTGATGCTGTGAGCCCAAGATTC 480  
Db 454 TGAAGAAATCCCACTGGGAGAGGGAAGAAATCAGCTGTGATGCTGTGAGCCCAAGATTC 513  
QY 481 ATAAGATAGCCACTCTGGGCTTTGGCAGAGCATTGGGACCTCTCCAGAAAGGCTTTACG 540  
Db 514 ATBAATAGCCACTCTGGGCTTTGGCAGAGCATTGGGACCTCTCCAGAAAGGCTTTACG 573  
QY 541 CAGAAATCTGTGTGTGACCTCTTTCGATGAACCTGCAAGAAAGGCTCTGAGAAAGAG 600  
Db 574 CAGAAATCTGTGTGTGACCTCTTTCGATGAACCTGCAAGAAAGGCTCTGAGAAAGAG 633  
QY 601 GGAATATTGTTTATTAACCACTTACATCACTCAAGAGCGGTGCAATCCGA 660  
Db 634 GGAATATTGTTTATTAACCACTTACATCACTCAAGAGCGGTGCAATCCGA 693  
QY 661 GCGAGGGGCGGTGGAAGCTGCGCAAGGTGGGGCTTTGGCATCTCAATTCGATCCGGG 720  
Db 694 GCGAGGGGCGGTGGAAGCTGCGCAAGGTGGGGCTTTGGCATCTCAATTCGATCCGGG 753  
QY 721 CCTCCTTCTCATCTACAGTCTCTCAACAGTATTCAGGAATACAGAGTGGCTGCCA 780  
Db 754 CCTCCTTCTCATCTACAGTCTCTCAACAGTATTCAGGAATACAGAGTGGCTGCCA 813  
QY 781 AATTCCAACAGCTGTATTAAGGTGGAAGATGCAAAATGATGTCCAGAAATGGCTTCTC 840  
Db 814 AATTCCAACAGCTGTATTAAGGTGGAAGATGCAAAATGATGTCCAGAAATGGCTTCTC 873  
QY 841 ATGGATCAAAATTTGTCATTCAGCTBAAGATGGGGCAAAAGCTTACAGTACTGAT 900  
Db 874 ATGGATCAAAATTTGTCATTCAGCTBAAGATGGGGCAAAAGCTTACAGTACTGAT 933  
QY 901 CCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGACAGGTGTACTGTCA 960  
Db 934 CCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGACAGGTGTACTGTCA 993  
QY 961 GTGACATCTGACAGCTGGGATGTTGGCAGGGGTGCCATGATATGCGGTGGAGCCT 1020  
Db 994 GTGACATCTGACAGCTGGGATGTTGGCAGGGGTGCCATGATATGCGGTGGAGCCT 1053  
QY 1021 TTATATCATGGAGACACTCTCATTTAAGATCTTGGGCTGGTCCAAAGAGACATC 1080  
Db 1054 TTATATCATGGAGACACTCTCATTTAAGATCTTGGGCTGGTCCAAAGAGACATC 1113  
QY 1081 TGCGGCTGTGCTGTGACTGACAGAAACAAGGTGAGTGTGCTTCCAGTATATC 1140

Db 1114 TGCGGCTGTGCTGTGACTGACAGAAACAAGGTGAGTGTGCTTCCAGTATATC 1173  
QY 1141 AGTTACCAAGSTAAATTTTCCAACTACAGCTGTGGATGAGATGCAAGCAAGCT 1200  
Db 1174 AGTTACCAAGSTAAATTTTCCAACTACAGCTGTGGATGAGATGCAAGCAAGCT 1233  
QY 1201 TCTTACCACTGGGCTCAATTCAGTGGCAGTGAAGAGCCAGAGCCATCATGAGAGG 1260  
Db 1234 TCTTACCACTGGGCTCAATTCAGTGGCAGTGAAGAGAGCCAGAGCCATCATGAGAGG 1293  
QY 1261 TTATGAGCTGTGCTGAGCCCTCAATATCACTGAGTCTGAGCCATGAGAGAGGACAG 1320  
Db 1294 TTATGAGCTGTGCTGAGCCCTCAATATCACTGAGTCTGAGCCATGAGAGAGGACAG 1353  
QY 1321 ACATCACTTTTGGATCCAGCTGAGAGTGGCTGAGCCAGCTCATGATGATATATCA 1380  
Db 1354 ACATCACTTTTGGATCCAGCTGAGAGTGGCTGAGCCAGCTCATGATGATATATCA 1413  
QY 1381 AGTATTTCTTCTCATCACTCCACAGAGACACATGACTGTATGATGATCAAGACA 1440  
Db 1414 AGTATTTCTTCTCATCACTCCACAGAGACACATGACTGTATGATGATCAAGACA 1473  
QY 1441 TGAATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1500  
Db 1474 TGAATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1533  
QY 1501 TGCTGCTGAGTCTGTAGAAACAGTAAAGAAAGAAAGTTCATGCTGTGCTGAGATC 1560  
Db 1534 TGCTGCTGAGTCTGTAGAAACAGTAAAGAAAGAAAGTTCATGCTGTGCTGAGATC 1593  
QY 1561 CTGGGCTGCAACTTGTGAAAACTCTCTTCACTAACAATTTGATCCATTCCTTCA 1620  
Db 1594 CTGGGCTGCAACTTGTGAAAACTCTCTTCACTAACAATTTGATCCATTCCTTCA 1653  
QY 1621 AAGCAAACTATTTATGATGCTTGTGATATCTTGTGATATCTTCAATTCCT 1680  
Db 1654 AAGCAAACTATTTATGATGCTTGTGATATCTTGTGATATCTTCAATTCCT 1713  
QY 1681 GATCTGAAAAAAGATTCATCTCCCTCCCTCCCAACATGATGATCAATATGTA 1740  
Db 1714 GATCTGAAAAAAGATTCATCTCCCTCCCTCCCAACATGATGATCAATATGTA 1773  
QY 1741 GGGATTCAGTGGGGGATTTCTTATATCACTCTTAAACAATGTTTCACTTTAA 1800  
Db 1774 GGGATTCAGTGGGGGATTTCTTATATCACTCTTAAACAATGTTTCACTTTAA 1833  
QY 1801 AGTAAACACTTAATTAATTTTGGAGATCTTGA 1835  
Db 1834 AGTAAACACTTAATTAATTTTGGAGATCTTGA 1868

RESULT 12  
AAK94491  
ID AAK94491 standard; cDNA; 1860 BP.  
XX  
AC AAK94491;  
XX  
DT 06-NOV-2001 (first entry)  
DE Human full-length cDNA, SEQ ID NO: 3328.  
XX  
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.  
OS Homo sapiens.  
XX  
PN EPI130094-A2.  
XX  
PD 05-SEP-2001.  
XX  
PF 07-JUL-2000; 2000EP-0114089.  
XX  
PR 08-JUL-1999; 99JP-0194486.



PR 11-JAN-2000; 2000JP-0118774.  
PR 02-MAY-2000; 2000JP-0183765.  
PA (HELI-) HELIX RES INST.  
PI Ota T, Nishikawa T, Isogai T, Hayaishi K, Ishii S, Kawai Y,  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H,  
XX WPI; 2001-524255/58.  
DR P-PSDB; AAM93559.  
XX  
XX  
PT 830 Primers useful for synthesizing full length cDNA clones and their  
PT use in genetic manipulation -  
PS Claim 8; SEQ ID NO 3328; 1380bp + sequence listing; English.  
XX  
XX The invention relates to primers for synthesizing full length cDNA  
XX clones. 830 cDNA molecules encoding a human protein have been  
XX isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
XX molecules have been determined. Primers for synthesizing the full length  
XX cDNA are useful for clarifying the function of the protein encoded by  
XX the cDNA. The full length clones were obtained by construction of full  
XX length enriched cDNA libraries that were synthesised by the oligo-capping  
XX method. The primers enable the production of the full length cDNA easily  
XX without any special methods. The present sequence is a full length  
XX human cDNA of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in CD-ROM format directly from EPO.  
XX  
XX Sequence 1860 BP; 515 A; 423 C; 451 G; 471 T; 0 other;  
SQ  
Query Match 98.3%; Score 1820.4; DB 22; Length 1860;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1824; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 GGTAGGCCGCGAGCTTATGCTCTGGAGCGCGCTCGCGCGCGCTCAGAGCGCCCTA 60  
DB 31 GGTAGGCCGCGAGCTTATGCTCTGGAGCGCGCTCGCGCGCGCTCAGAGCGCCCTA 90  
QY 61 TCAATATATCTTAACAAGAAACCAACCTGGAAGAAATGAATTCCTTATCTTCGAT 120  
DB 91 TCAATATATCTTAACAAGAAACCAACCTGGAAGAAATGAATTCCTTATCTTCGAT 150  
QY 121 TTTTCGGTGTGTTACCTTTATCCCTGTCTGTGGAAAGCTATATGCAAGATGCA 180  
DB 151 TTTTCGGTGTGTTACCTTTATCCCTGTCTGTGGAAAGCTATATGCAAGATGCA 210  
QY 181 TCTTAAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTTGCTA 240  
DB 211 TCTTAAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTTGCTA 270  
QY 241 AAGCAATCATCACTAGCTGTTATGTATTAAGCCAGAAAGATCTATAGAGATGG 300  
DB 271 AAGCAATCATCACTAGCTGTTATGTATTAAGCCAGAAAGATCTATAGAGATGG 330  
QY 301 CACTTCGTGTGATACGTGTGAGCCAGAGCTGAGGTGCTCAAGAACTTAGAAAAAGCA 360  
DB 331 CACTTCGTGTGATACGTGTGAGCCAGAGCTGAGGTGCTCAAGAACTTAGAAAAAGCA 390  
QY 361 TCCAAATTTATGACCAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGAGCCAG 420  
DB 391 TCCAAATTTATGACCAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGAGCCAG 450  
QY 421 TGAAGATATCCCACTGGAGAGGGAGAAAGATCACTGTATGTGTGAGCCAGAAATTC 480  
DB 451 TGAAGATATCCCACTGGAGAGGGAGAAAGATCACTGTATGTGTGAGCCAGAAATTC 510  
QY 481 ATAAGATGCAATCCTGGGTCTGTGAGCAGACATTTGGGACTCTCCAGAAAGCAATTAC 540  
DB 511 ATAAGATGCAATCCTGGGTCTGTGAGCAGACATTTGGGACTCTCCAGAAAGCAATTAC 570  
QY 541 CAGAAGTTCTGGTGTGACCTCTTTCATGAACTGACAGAGAGGGCTCAGAAAGCAAG 600

DB 571 CAGAAGTTCTGGTGTGACCTCTTTCATGAACTGACAGAGAGGGCTCAGAAAGCAAG 630  
QY 601 GAAAGATTTGTTGTTATTAACCAACTTATCAATCACTCAAGAGCGGTCAATACGAA 660  
DB 631 GAAAGATTTGTTGTTATTAACCAACTTATCAATCACTCAAGAGCGGTCAATACGAA 690  
QY 661 CGCAGGGGGCGGTGAAGCTGCCAAGGTGGGGGCTTTGGCATCTCTCATTCGATCCGTGG 720  
DB 691 CGCAGGGGGCGGTGAAGCTGCCAAGGTGGGGGCTTTGGCATCTCTCATTCGATCCGTGG 750  
QY 721 CTTCTTCTCATCTTACAGTCTTCAACAGAGTATTCAGAAATACAGAGATGGCGTCCA 780  
DB 751 CTTCTTCTCATCTTACAGTCTTCAACAGAGTATTCAGAAATACAGAGATGGCGTCCA 810  
QY 781 AAATTCAGAGCGCTGATTAAGGTGGAATGAGAAATGATGTCAAGATGGCTCTC 840  
DB 811 AGATTCAGAGCGCTGATTAAGGTGGAATGAGAAATGATGTCAAGATGGCTCTC 870  
QY 841 ATGGATCAAAATTTGTCATTCAGCTAAAGTGGGGGCAAGACCTACAGATTAAGT 900  
DB 871 ATGGATCAAAATTTGTCATTCAGCTAAAGTGGGGGCAAGACCTACAGATTAAGT 930  
QY 901 CTTTCAACTGTGACAGAGATCACTGGAGCAAAATATCCAGAACAGTTGTACTGTCA 960  
DB 931 CTTTCAACTGTGACAGAGATCACTGGAGCAAAATATCCAGAACAGTTGTACTGTCA 990  
QY 961 GTGACATCTGACAGCTGGAGATTTGGGAGGTGTCAGATGATGGGTGGAGCCT 1020  
DB 991 GTGACATCTGACAGCTGGAGATTTGGGAGGTGTCAGATGATGGGTGGAGCCT 1050  
QY 1021 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTC 1080  
DB 1051 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTC 1110  
QY 1081 TGGGCTGTGTCTGTGACCTGACAGAGAACCAAGGTGAGTGTGCTTCCAGTATATC 1140  
DB 1111 TGGGCTGTGTCTGTGACCTGACAGAGAACCAAGGTGAGTGTGCTTCCAGTATATC 1170  
QY 1141 AGTTACCAAGGTAAATTTTCAACATCAAGCTGTGATGAGATCTGACGAGGAACT 1200  
DB 1171 AGTTACCAAGGTAAATTTTCAACATCAAGCTGTGATGAGATCTGACGAGGAACT 1230  
QY 1201 TCTTACCACTGGGCTCAATTCACCTGCGAGTGAAGAGCCAGGGCCATATGAGAGAG 1260  
DB 1231 TCTTACCACTGGGCTCAATTCACCTGCGAGTGAAGAGCCAGGGCCATATGAGAGAG 1290  
QY 1261 TTATGAGCTGTGACAGCCCTCAATATCACTCAAGTCTGAGCCATGAGAAAGAGCAG 1320  
DB 1291 TTATGAGCTGTGACAGCCCTCAATATCACTCAAGTCTGAGCCATGAGAAAGAGCAG 1350  
QY 1321 ACATCACTTTTGGATTCAGAGTGAATGCTGTGAGAGCACTTATGATGACTTATACA 1380  
DB 1351 ACATCACTTTTGGATTCAGAGTGAATGCTGTGAGAGCACTTATGATGACTTATACA 1410  
QY 1381 AGTATTCCTTCTTCATCACTCCAGAGGAGACACATGATGATGATCCAAAGCAGA 1440  
DB 1411 AGTATTCCTTCTTCATCACTCCAGAGGAGACACATGATGATGATCCAAAGCAGA 1470  
QY 1441 TGAATGTTGCTGCTGTTTGGGCTGTTTCTTATGTTGTTGACAGATGGAAGAAA 1500  
DB 1471 TGAATGTTGCTGCTGTTTGGGCTGTTTCTTATGTTGTTGACAGATGGAAGAAA 1530  
QY 1501 TGTGCTAGTCTTAAAGACAGTGAAGAAAGAGTTTTCATGCTTGTGCGCAGAAATC 1560  
DB 1531 TGTGCTAGTCTTAAAGACAGTGAAGAAAGAGTTTTCATGCTTGTGCGCAGAAATC 1590  
QY 1561 CTGGGCTGCAACTTTGAAAGACCTCTTCACTAATTAATTCATCAATTCATCTCA 1620  
DB 1591 CTGGGCTGCAACTTTGAAAGACCTCTTCACTAATTAATTCATCAATTCATCTCA 1650  
QY 1621 AAGCAACTTATTTGATGCTTGTGTTATTTATCTTGTGATCTTCAATTTCTCT 1680  
DB 1651 AAGCAACTTATTTGATGCTTGTGTTATTTATCTTGTGATCTTCAATTTCTCT 1710

QY 1681 GATTCTAGAAAAAGAAATCATTCCTCCCTCCACACATAGAAATCAATATGTA 1740  
DB 1711 GATTCTAGAAAAAGAAATCATTCCTCCCTCCACACATAGAAATCAATATGTA 1770  
QY 1741 GGGATTACAGTGGGGGCAATTTCTTTATATACCTCTTAAAAACATTTTCCACTTTAA 1800  
DB 1771 GGGATTACAGTGGGGGCAATTTCTTTATATACCTCTTAAAAACATTTTCCACTTTAA 1830  
QY 1801 AGTAAACCTTATATATTTTGGAGATC 1830  
DB 1831 AGTAAACCTTATATATTTTGGAGATC 1860

RESULT 13  
AAA40493  
ID AAA40493 standard; cDNA; 1767 BP.  
XX AAA40493;  
AC  
XX  
XX  
DT 16-NOV-2000 (first entry)  
XX  
DE Human fetal kidney cDNA fragment AM282\_11.  
XX  
XX Secreted protein; cytosolic; immunostimulatory; antimicrobial;  
XX antiviral; immunosuppressive; antiinflammatory; vulnerary; cytokine;  
XX cell proliferation; differentiation; regulation; treatment; tumor;  
XX autoimmune disease; inflammatory disorder; wound; microbial infection;  
XX viral disease; graft versus host reaction suppression; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO20037630-A1.  
XX  
XX 29-JUN-2000.  
XX  
XX 22-DEC-1999; 99WO-US31005.  
XX  
XX 23-DEC-1998; 98US-0220876.  
XX  
XX (GENY) GENETICS INST INC.  
XX  
XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
XX Merberg D, Treacy M, Bowman MR;  
XX  
XX MPI: 2000-442661/38.  
XX  
XX P-PSDB; AAB10229.  
XX  
XX Secreted human proteins AS296-11 and AS34-11, useful for treating  
XX tumors, autoimmune diseases, inflammatory disorders, wounds, microbial  
XX infections and viral diseases -  
XX  
XX  
XX Disclosure; Page 198; 293pp; English.  
XX  
XX This invention describes novel secreted human proteins (I) which have  
XX cytostatic, immunostimulatory, antimicrobial, antiviral,  
XX immunosuppressive, antiinflammatory and vulnerary activity and which act  
XX as cytokine, cell proliferation or differentiation regulators. (I)  
XX is useful for treating tumors, autoimmune diseases, inflammatory  
XX disorders, wounds, microbial infections and viral diseases. (I) is also  
XX useful for suppressing graft versus host reaction. AAA40490-A40580  
XX represent cDNA fragments that encode the secreted proteins  
XX AAB10226-810288 described in the method of the invention.  
XX  
XX  
XX Sequence 1767 BP; 512 A; 384 C; 418 G; 453 T; 0 other;  
XX  
XX  
XX Query Match 94.6%; Score 1751.2; DB 21; Length 1767;  
XX Best Local Similarity 99.8%; Pred. No. 0;  
XX Matches 1753; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 96 AAAATGAATTCCTTATCTTGCATTTTGGTGTGTTCACCTTTTATCCCTGTGCTCT 155  
DB 3 AAGATGAATTCCTTATCTTGCATTTTGGTGTGTTCACCTTTTATCCCTGTGCTCT 62

QY 156 GGAAGAGTATATGCAAGATGSCATCTTAAGAGACTTTTGAAGAAATTAAGAGAA 215  
DB 63 GGAAGAGTATATGCAAGATGSCATCTTAAGAGACTTTTGAAGAAATTAAGAGAA 122  
QY 216 ATAGCCAGCTGTGAGATGTGTGTTAAAGCAATCATCAACCTAGCTTTATGTAAAGCC 275  
DB 123 ATAGCCAGCTGTGAGATGTGTGTTAAAGCAATCATCAACCTAGCTTTATGTAAAGCC 182  
QY 276 CAGAACGATCTTATGAGCCGATGAGCACTTGTGTATATCTGTGAACCCAGCTAGT 335  
DB 183 CAGAACGATCTTATGAGCCGATGAGCACTTGTGTATATCTGTGAACCCAGCTAGT 242  
QY 336 GGCCTCAAGAACCTAGAAAAAGCCATCCAAATTATGTACCAAAACCTGACAGCAAGATGG 395  
DB 243 GGCCTCAAGAACCTAGAAAAAGCCATCCAAATTATGTACCAAAACCTGACAGCAAGATGG 302  
QY 396 CTGAGAAAGTTTCACTGAGACCGAGTAAGTAATCCCACTGGGAGAGGGAGAAATCA 455  
DB 303 CTGAGAAAGTTTCACTGAGACCGAGTAAGTAATCCCACTGGGAGAGGGAGAAATCA 362  
QY 456 GCTGTGATGCTGAGCCCAAGAAATTCATTAATATCCATCTCTGGGTCTTGGACAGACATT 515  
DB 363 GCTGTGATGCTGAGCCCAAGAAATTCATTAATATCCATCTCTGGGTCTTGGACAGACATT 422  
QY 516 GGGACTCCTCCAGAAAGGCAATTAACAGAGAGTTCTGGTGTGAGCTCTTTGATGAATCTG 575  
DB 423 GGGACTCCTCCAGAAAGGCAATTAACAGAGAGTTCTGGTGTGAGCTCTTTGATGAATCTG 482  
QY 576 CAGAGAAAGGCTCAGAAAGCAAGAGGAAAGTTGTATTAAACCAACCTTACATCAAC 635  
DB 483 CAGAGAAAGGCTCAGAAAGCAAGAGGAAAGTTGTATTAAACCAACCTTACATCAAC 542  
QY 636 TACTCAAGAGAGGTGCAATACCGAACCGAGGGGCGGTGAGAGCTGCCAAGGTGGGGCT 695  
DB 543 TACTCAAGAGAGGTGCAATACCGAACCGAGGGGCGGTGAGAGCTGCCAAGGTGGGGCT 602  
QY 696 TTGGCATCTCTCATTCGATCCGTGGCTCTCTTCATCTTCAAGTCCCTACACAGATATT 755  
DB 603 TTGGCATCTCTCATTCGATCCGTGGCTCTCTTCATCTTCAAGTCCCTACACAGATATT 662  
QY 756 CAGAAATPACAGATGCGTGGTCCCAAAATTCACACAGCTGTATTAAGTGAAGATGA 815  
DB 663 CAGAAATPACAGATGCGTGGTCCCAAAATTCACACAGCTGTATTAAGTGAAGATGA 722  
QY 816 GAATGATGTAAAGATGCTTCTCATGAGATCAAAATTTGATTCAGCTTAAAGATGGG 875  
DB 723 GAATGATGTAAAGATGCTTCTCATGAGATCAAAATTTGATTCAGCTTAAAGATGGG 782  
QY 876 GCAAGAGCTTACAGATGATGATTCCTTCAACACTGTAGAGAGATCACTGGAGCAAA 935  
DB 783 GCAAGAGCTTACAGATGATGATTCCTTCAACACTGTAGAGAGATCACTGGAGCAAA 842  
QY 936 TATCAGAAACAGTTGTATCTGTCAGTGAATCTGACAGCTGGAGATGTTGGCAGGGT 995  
DB 843 TATCAGAAACAGTTGTATCTGTCAGTGAATCTGACAGCTGGAGATGTTGGCAGGGT 902  
QY 996 GGCATGATGATGCGGTGAGGCTTTATATCATGGAAGACATCTCACTTTAATGAAT 1055  
DB 903 GGCATGATGATGCGGTGAGGCTTTATATCATGGAAGACATCTCACTTTAATGAAT 962  
QY 1056 CTTGGGCTGCGTCCAAAGAGAGACTCTGCGGTGGTGTCTGAGACTGGAGAAACAGAGT 1115  
DB 963 CTTGGGCTGCGTCCAAAGAGAGACTCTGCGGTGGTGTCTGAGACTGGAGAAACAGAGT 1022  
QY 1116 GGAATGATGCTTCAAGATATATATCAATTAACAGAGTAATATTTTCAACTACAGTCTG 1175  
DB 1023 GGAATGATGCTTCAAGATATATATCAATTAACAGAGTAATATTTTCAACTACAGTCTG 1082  
QY 1176 GTGATGAGATGTAGCAGAGAACTTTTACCCACTGGGCTGCAATTCACCTGGAGATGA 1235  
DB 1083 GTGATGAGATGTAGCAGAGAACTTTTACCCACTGGGCTGCAATTCACCTGGAGATGA 1142

QY 1236 AAGCCAGGCGCATCATGAGAGGATTATGAGCTGCTGAGAGCCCTCAATATCATCTCAG 1295  
 DB 1143 AAGCCAGGCGCATCATGAGAGGATTATGAGCTGCTGAGAGCCCTCAATATCATCTCAG 1292  
 QY 1296 GTCTGAGCCATGAGAGAGGACAGACATCACTTTTGGATCCAACTGAGTCCCTGGA 1355  
 DB 1203 GTCTGAGCCATGAGAGAGGACAGACATCACTTTTGGATCCAACTGAGTCCCTGGA 1262  
 QY 1356 GCCAGTCTGATGATGATTAACAAGATTCTCTTCCATCATCTCCACGAGACACC 1415  
 DB 1263 GCCAGTCTGATGATGATTAACAAGATTCTCTTCCATCATCTCCACGAGACACC 1322  
 QY 1416 ATGACTGTCATGATCCAAAGACATGATGATGCTGCTGCTGCTGCTGCTGCTGCT 1475  
 DB 1323 ATGACTGTCATGATCCAAAGACATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1382  
 QY 1476 TATGTTGTCAGATGATGAGAAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1535  
 DB 1383 TATGTTGTCAGATGATGAGAAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1442  
 QY 1536 GTTTTCATGCTTCTGCGCAGGAATCTGCGTCTGCACTTTGGAAAACCTCCTTCACAT 1595  
 DB 1443 GTTTTCATGCTTCTGCGCAGGAATCTGCGTCTGCACTTTGGAAAACCTCCTTCACAT 1502  
 QY 1596 AACAAATTCATCCAAATTCATCTTCAAGACCACTATTCATGCTTCTGCTGCTGCTGCT 1655  
 DB 1503 AACAAATTCATCCAAATTCATCTTCAAGACCACTATTCATGCTTCTGCTGCTGCTGCT 1562  
 QY 1656 TTTCTTGATATCTTCCAAATTCCTGATCTGATGCTGATGCTGATGCTGATGCTGATGCT 1715  
 DB 1563 TTTCTTGATATCTTCCAAATTCCTGATCTGATGCTGATGCTGATGCTGATGCTGATGCT 1622  
 QY 1716 CACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1775  
 DB 1623 CACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1682  
 QY 1776 TTAATAACATGCTTCCATCTTAAAGTAACACATTAATAATTTTGAAGATCTCTGA 1835  
 DB 1683 TTAATAACATGCTTCCATCTTAAAGTAACACATTAATAATTTTGAAGATCTCTGA 1742  
 QY 1836 AAAAAAAAAAAAAAAAAA 1851  
 DB 1743 AAAAAAAAAAAAAAAAAA 1758

RESULT 14  
 AAV02296 ID AAV02296 standard; DNA; 1778 BP.  
 XX AAV02296;  
 AC  
 XX 21-MAY-1998 (first entry)  
 DT  
 XX Human secreted protein AM282 full-length cDNA clone.  
 DE  
 XX Secreted protein; AM282; cytokine; human; ds.  
 KM  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH CDS 17..1435  
 FT /\*tag= a  
 FT sig\_peptide 17..88  
 FT /\*tag= b  
 FT mat\_peptide 89..1432  
 FT /\*tag= b  
 XX  
 XX MO9739030-A2.  
 XX 23-OCT-1997.  
 PD  
 XX 16-APR-1997; 97WO-US06475.  
 PF  
 XX

PR 13-JAN-1997; 97US-0783520.  
 PR 18-APR-1996; 96US-0634325.  
 XX  
 PA (GENE) GENETICS INST INC.  
 XX  
 PI Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA;  
 PI Spaulding V;  
 XX  
 DR WPI: 1997-526400/48.  
 DR P-PSDB: AAM33604.  
 XX  
 XX New isolated secretory proteins AM340, AM282 and AK583 - possibly  
 PT have cytokine, cell proliferation/differentiation regulating,  
 PT immunomodulating activities, etc.  
 PT  
 XX  
 PS Claim 15; Page 44-45; 59pp; English.  
 XX  
 CC This cDNA clone encodes a protein (see W33604) designated AM282.  
 CC It was identified as "y195b10.r1 human EST 30142.5" (GenBank  
 CC accession No. R77830) in a database search using a partial AM282  
 CC clone (see T97398) obtained from a human foetal kidney cDNA  
 CC library using methods selective for cDNAs encoding secreted  
 CC proteins. AM282 is deposited in ATCC 98026 together with clones  
 CC AM340 (see T97397) and AK583 (see V02297), which are also claimed.  
 CC AM282 protein can be used in a claimed method for preventing,  
 CC treating or ameliorating a medical condition. It may exhibit  
 CC cytokine, cell proliferation (either inducing or inhibiting) or  
 CC cell differentiation (either inducing or inhibiting) activity or  
 CC may induce production of other cytokines in certain cell  
 CC populations. It may also exhibit e.g. immune stimulating or  
 CC suppressing activity, haematopoiesis regulating activity, tissue  
 CC growth activity, activin/inhibin activity, chemotactic or  
 CC chemokinetic activity, haemostatic or thrombolytic activity,  
 CC receptor/ligand activity, anti-inflammatory activity, tumour  
 CC inhibition activity, or other activities. No evidence of any of  
 CC these activities is given in the specification.  
 CC  
 XX  
 SO Sequence 1778 BP; 514 A; 386 C; 422 G; 456 T; 0 other;

Query Match 94.6%; Score 1751; DB 18; Length 1778;  
 Best Local Similarity 99.4%; Pred. No. 0;  
 Matches 1757; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 85 AACTGAAAAAAAAAATGAAATCTCTATCTTCGATTTTCGGGAGTTCACCTTTAT 144  
 DB 3 AGCTTGGCAGAGATGAAATCTCTATCTTCGATTTTCGGGAGTTCACCTTTAT 62  
 QY 145 CCTGTGCTCTGGGAAAGCTATATGCAAGATGGCATCTTAAGAGACTTTTGAAGAA 204  
 DB 63 CCTGTGCTCTGGGAAAGCTATATGCAAGATGGCATCTTAAGAGACTTTTGAAGAA 122  
 QY 205 TAAAGAAAGAAATGCAAGCTGTGAGATGTGTCTAAAGCATCATCACTAGCTGTT 264  
 DB 123 TAAAGAAAGAAATGCAAGCTGTGAGATGTGTCTAAAGCATCATCACTAGCTGTT 182  
 QY 265 ATGCTAAGCCCAAGACAGATCCATGAGAGATTTGGACTTCGTGATAGTTGGAC 324  
 DB 183 ATGCTAAGCCCAAGACAGATCCATGAGAGATTTGGACTTCGTGATAGTTGGAC 242  
 QY 325 CCAAGCTGATGCTCCAGAACTAGAAAAAGCAATCAATATTTATACAAAACCTGC 384  
 DB 243 CCAAGCTGATGCTCCAGAACTAGAAAAAGCAATCAATATTTATACAAAACCTGC 302  
 QY 385 AGCAAGATGGCTGAGAAAGTTCACTGAGACCAAGTGAATGCCCATCTGGAGAGGG 444  
 DB 303 AGCAAGATGGCTGAGAAAGTTCACTGAGACCAAGTGAATGCCCATCTGGAGAGGG 362  
 QY 445 GAGAAGATGACGTGATGCTGAGCCAGAAATTCATTAAGATAGCATCTGGGCTTG 504  
 DB 363 GAGAAGATGACGTGATGCTGAGCCAGAAATTCATTAAGATAGCATCTGGGCTTG 422  
 QY 505 GCACAGCATTTGGAGACTCTCCAGAAAGCATTCACAGAGAAATTTCTGTGTGACTCTT 564

```

Db 423 GCACGACGATTGGGACCTCTCCAGAAAGGCTTACAGAGAAGTTCTGGTGGACCTCTT 482
Qy 565 TCGATGAACTGCAAGAAAGGGCTTCAGAAAGCAAGGGAAATGTTTATTAACCAAC 624
Db 483 TCGATGAACTGCAAGAAAGGGCTTCAGAAAGCAAGGGAAATGTTTATTAACCAAC 542
Qy 625 CTTTACATCACTACTCAAGAGCGGTGCAATACCGAAGCGAGGGCGGTGGAAAGCTGCA 684
Db 543 CTTTACATCACTACTCAAGAGCGGTGCAATACCGAAGCGAGGGCGGTGGAAAGCTGCA 602
Qy 685 AGGTGGGGGCTTTGGCATCTCTCATTCGATCCGTGGCTCTCTTCATCTCACTGCTCTC 744
Db 603 AGGTGGGGGCTTTGGCATCTCTCATTCGATCCGTGGCTCTCTTCATCTCACTGCTCTC 662
Qy 745 ACACAGGATTCAGGAATACAGAGATGGCGGCCCAAAATTCACACCGTATTAAGG 804
Db 663 ACACAGGATTCAGGAATACAGAGATGGCGGCCCAAAATTCACACCGTATTAAGG 722
Qy 805 TGAAGATGCAAGAAATGATGTCAGAAATGGCTTCATGGGATCAAAATTTGTCATTCAGC 864
Db 723 TGAAGATGCAAGAAATGATGTCAGAAATGGCTTCATGGGATCAAAATTTGTCATTCAGC 782
Qy 865 TAAAGATGGGGCAAAAGACTTACCCAGATGATTCCTTCAACCTGTAGCAGATCA 924
Db 783 TAAAGATGGGGCAAAAGACTTACCCAGATGATTCCTTCAACCTGTAGCAGATCA 842
Qy 925 CTGGAGCAAAATATCCAGAACGTTTACTGCTGATGGATGACATTTGGACAGCTGGGATG 984
Db 843 CTGGAGCAAAATATCCAGAACGTTTACTGCTGATGGATGACATTTGGACAGCTGGGATG 902
Qy 985 TTGGGCGAGGTGCCATGATGATGCGGTGGAGCGCTTTATATCATGGGAAGCACTCTCAC 1044
Db 903 TTGGGCGAGGTGCCATGATGATGCGGTGGAGCGCTTTATATCATGGGAAGCACTCTCAC 962
Qy 1045 TTATTTAAAGATCTTGGGCTGCTGCCAAAGAGACTCTGCGGCTGGTGTCTGAGCTGCAG 1104
Db 963 TTATTTAAAGATCTTGGGCTGCTGCCAAAGAGACTCTGCGGCTGGTGTCTGAGCTGCAG 1022
Qy 1105 AAGAACAGGTGGAATTTGGTGTCTCCGATATTAACATTAACAAAGGTAATATTTTCCA 1164
Db 1023 AAGAACAGGTGGAATTTGGTGTCTCCGATATTAACATTAACAAAGGTAATATTTTCCA 1082
Qy 1165 ACTACAGCTGGGTGATGAGTGTGACGAGAGAACTTCTTACCCACTGGGCTGCATTTCA 1224
Db 1083 ACTACAGCTGGGTGATGAGTGTGACGAGAGAACTTCTTACCCACTGGGCTGCATTTCA 1142
Qy 1225 CTGGCAGTGAAGAGGCCAGGGCCATCATGAGAGAGGTTATGAGCTGTGACGCCCTCA 1284
Db 1143 CTGGCAGTGAAGAGGCCAGGGCCATCATGAGAGAGGTTATGAGCTGTGACGCCCTCA 1202
Qy 1285 ATATCACTCAGGTCTGTAGCCCATGGAAGAGGACAGACATCACTTTTGGATCCAAAGCTG 1344
Db 1203 ATATCACTCAGGTCTGTAGCCCATGGAAGAGGACAGACATCACTTTTGGATCCAAAGCTG 1262
Qy 1345 GAGTGCCTGGAGCCAGTCTACTTATGATGATTAACAAGATTTCTTCTTCATCACTGCC 1404
Db 1263 GAGTGCCTGGAGCCAGTCTACTTATGATGATTAACAAGATTTCTTCTTCATCACTGCC 1322
Qy 1405 ACGGAGACCACTGACTGTATGATCCAAAGCAGATGAATGTTGCTGCTGTTTGGG 1464
Db 1323 ACGGAGACCACTGACTGTATGATCCAAAGCAGATGAATGTTGCTGCTGTTTGGG 1382
Qy 1465 CTGTTGTTTCTTATGTTGTGACACATGGAAGAAATGCTGCTAGGTTCTTGAAGAAAGCT 1524
Db 1383 CTGTTGTTTCTTATGTTGTGACACATGGAAGAAATGCTGCTAGGTTCTTGAAGAAAGCT 1442
Qy 1525 AAGAAAGAAAGCTTTTCACTGCTTGGGCGAGGAATCTGGGCTGCAACTTTTGAAGAACT 1584
Db 1443 AAGAAAGAAAGCTTTTCACTGCTTGGGCGAGGAATCTGGGCTGCAACTTTTGAAGAACT 1502
Qy 1585 CCTCTTCACTAATCAATTTCAATTCATCTTCAAGCAACACTCTATTTCACTGCTTT 1644
Db 1503 CCTCTTCACTAATTTCAATTCATTCATCTTCAAGCAACACTCTATTTCACTGCTTT 1562

```

```

Qy 1645 CTGTTATTTATCTTCTTGATACCTTCCAAATTTCTGATTTCTAGAAAAAGAAATCATTTCT 1704
Db 1563 CTGTTATTTATCTTCTTGATACCTTCCAAATTTCTGATTTCTAGAAAAAGAAATCATTTCT 1622
Qy 1705 CCCCTCCCTCCACACATAGAAATCAACATATGTTAGGATTTACGTGGGGGCAATTTCTT 1764
Db 1623 CCCCTCCCTCCACACATAGAAATCAACATATGTTAGGATTTACGTGGGGGCAATTTCTT 1682
Qy 1765 TATATCACTCTTAAACATTTGTTTCCATTTTAAAGTTAAACCTTAATTAATTTTGG 1824
Db 1683 TATATCACTCTTAAACATTTGTTTCCATTTTAAAGTTAAACCTTAATTAATTTTGG 1742
Qy 1825 AAGATCTCTGAAAAAATTTTAAAAA 1851
Db 1743 AAGATCTCTGAAAAAATTTTAAAAA 1769

RESULT 15
AAH9703
ID AAH9703 standard; cDNA; 1895 BP.
XX
AC AAH9703;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human protein encoding cDNA sequence SEQ ID NO:538.
XX
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; vitruide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnery; antilucer; osteopathic; eczema;
KW dermatological; antiallergic; antisthmatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiParkinsonian; infection;
KW immunostimulant; gene therapy; antilease therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder; ss.
XX
OS Homo sapiens.
XX
PN W0200153455-A2.
XX
PD 26-JUL-2001.
XX
PE 22-DEC-2000; 2000WO-US35017.
XX
PR 23-DEC-1999; 99US-0471275.
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
XX
PA (HVSF-) HVSQ INC.
XX
PI Tang YT, Liu C, Dmanac RT;
XX
DR WPI; 2001-457603/49.
XX
DR P-PSDB; AAM25762.
XX
PT Isolated human polynucleotides encoding polypeptides, useful for the
XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
PS Claim 1; Page 591; 1217pp; English.
XX
CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAH25963. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; vitruide; anti-HIV; fungicide; antimutagen;

```

CC cardiovascular; antihaemic; antiagregant; haemostatic; vulnery;  
CC antiulcer; osteoplastic; dermatological; antiallergic; antiaethmatic;  
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;  
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
encoding them can be used in gene therapy, antisense therapy and vaccine  
production. The proteins and polynucleotides are useful for screening for  
agents or antagonists of a protein and for the treatment and diagnosis  
of disorders associated with the activity of a protein e.g. inflammation,  
rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
infections, autoimmunity, genetic diseases, haematopoietic disorders,  
anaemia, platelet disorders, chromocytopenia, wounds, burns, ulcers,  
osteoporosis, severe combined immunodeficiency, eczema, allergic  
rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
CC neurological disorders.

XX Sequence 1895 BP, 530 A; 439 C; 450 G; 476 T; 0 other;

Query Match 87.4%; Score 1617.2; DB 22; Length 1895;  
Best Local Similarity 96.9%; Pred. No. 0;  
Matches 1811; Conservative 0; Mismatches 33; Indels 24; Gaps 15;

QY 1 GGCTAGGCGCGAGCTTATGCTCTGGAGCGCGCTTCGCGCGCTGAGAGCGCCCTTA 60  
DB 28 GGCTAGGCGCGAGCTTATGCTCTGGAGCGCGCTTCGCGCGCTGAGAGCGCCCTTA 87  
QY 61 TCAGATTATCTTAACAAGAAACCACTGGAAGAAAAAATGAATTCCTTCTTCGAT 120  
DB 88 TCAGATTATCTTAACAAGAAACCACTGGAAGAAAAAATGAATTCCTTCTTCGAT 147  
QY 121 TTTTCGGTGGTTCACCTTTATCCCTGTGCTGGGAAAGTATATGAAAGATGCA 180  
DB 148 TTTTCGGTGGTTCACCTTTATCCCTGTGCTGGGAAAGTATATGAAAGATGCA 207  
QY 181 TCTCTAAGAGAGCTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTGCTA 240  
DB 208 TCTCTAAGAGAGCTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTGCTA 267  
QY 241 AAGGAATCATCAACTGAGCTTTATGTTAAGCCCAAGACAGATCCATAGAGATGCG 300  
DB 268 AAGGAATCATCAACTGAGCTTTATGTTAAGCCCAAGACAGATCCATAGAGATGCG 327  
QY 301 CACTTCGTGTGATGACTGTGAGCCAGACTGAGTGGCTCCAAGAACTTAGAAAAAGCA 360  
DB 328 CACTTCGTGTGATGACTGTGAGCCAGACTGAGTGGCTCCAAGAACTTAGAAAAAGCA 387  
QY 361 TCCAATTTATGTACCAAAACCTGACAGCAAGATGGGCTGAGAAAGTTCACTGAGCCAG 420  
DB 388 TCCAATTTATGTACCAAAACCTGACAGCAAGATGGGCTGAGAAAGTTCACTGAGCCAG 447  
QY 421 TGAAGATATCCCACTGGGAGAGGGAGAAATACGCTGTGATGTGAGCCAAAGATTC 480  
DB 448 TGAAGATATCCCACTGGGAGAGGGAGAAATACGCTGTGATGTGAGCCAAAGATTC 507  
QY 481 ATAAGATATGCACTCTGGGCTGTGGAGAGCAATTTGGGACTCTCCAGAAAGCAATTAG 540  
DB 508 ATAAGATATGCACTCTGGGCTGTGGAGAGCAATTTGGGACTCTCCAGAAAGCAATTAG 567  
QY 541 CAGAAATTTCTGGTGTGACTCTTTTCGATGAACTGCAAGAGAGGGCTCAGAAAGCAGAG 600  
DB 568 CAGAAATTTCTGGTGTGACTCTTTTCGATGAACTGCAAGAGAGGGCTCAGAAAGCAGAG 627  
QY 601 GGAAGATTTGTTTATTAACCACTTACATCACTCAAGGACGGTGTGAATACCGAA 660  
DB 628 GGAAGATTTGTTTATTAACCACTTACATCACTCAAGGACGGTGTGAATACCGAA 687  
QY 661 CGCAGGGGGGGTGGAGAGCTGCAAGGAGGGGGCTTTGGCATCTCTCATTTGGATCCGGG 720  
DB 688 CGCAGGGGGGGTGGAGAGCTGCAAGGAGGGGGCTTTGGCATCTCTCATTTGGATCCGGG 747  
QY 721 CCTCTCTCTCATCTACAGTCTCTCAACAGGATTTCAAGAAATACAGAGATGGCGTCCCA 780

DB 748 CCTCTCTCTCATCTACAGTCTCTCAACAGGATTTCAAGAAATACAGAGATGGCGTCCCA 807  
QY 781 AAATTCACAGAGCTGTATTAAGGTGAGAGATGAGCAAGATGATGCAAGATGGCTCTC 840  
DB 808 AGATTCACAGAGCTGTATTAAGGTGAGAGATGAGCAAGATGATGCAAGATGGCTCTC 867  
QY 841 ATGGGATCAAAATTTGTCAATTCAGTAAAGATGGGGGCAAGACCTTACCCAGATCTGAT 900  
DB 868 ATGGGATCAAAATTTGTCAATTCAGTAAAGATGGGGGCAAGACCTTACCCAGATCTGAT 927  
QY 901 CTTTCAACACTGTATGAGAGATCACTGGAGCAAAATTTCCAGAAACAGTTGTACTGTCA 960  
DB 928 CTTTCAACACTGTATGAGAGATCACTGGAGCAAAATTTCCAGAAACAGTTGTACTGTCA 987  
QY 961 GTGAGACATGGAAGAGTGGGAGTGGGAGGGTGCATGAGATGAGGGGTGAGAGCT 1020  
DB 988 GTGAGACATGGAAGAGTGGGAGTGGGAGGGTGCATGAGATGAGGGGTGAGAGCT 1047  
QY 1021 TTATATCATGGGAGAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTC 1080  
DB 1048 TTATATCATGGGAGAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTC 1107  
QY 1081 TGGCGTGTGTCTGTGAGACTGACAGAGAACAGAGTGGTGTGCTTCCAGTATATC 1140  
DB 1108 TGGCGTGTGTCTGTGAGACTGACAGAGAACAGAGTGGTGTGCTTCCAGTATATC 1167  
QY 1141 AGTTACAAGAGTAAATTTTCCAACTACAGTCTGTGATGAGATCTGACGACAGAACT 1200  
DB 1168 AGTTACAAGAGTAAATTTTCCAACTACAGTCTGTGATGAGATCTGACGACAGAACT 1227  
QY 1201 TCTTACCACTGGGCTCAATATCACTGAGCAAGAAAGCCAGAGGCTCATGAGAGAG 1260  
DB 1228 TCTTACCACTGGGCTCAATATCACTGAGCAAGAAAGCCAGAGGCTCATGAGAGAG 1287  
QY 1261 TTATGAGCTGTGACAGCCCTCAATATCACTGAGTCTGAGCCATGAGAGAGAGAG 1320  
DB 1288 TTATGAGCTGTGACAGCCCTCAATATCACTGAGTCTGAGCCATGAGAGAGAGAG 1347  
QY 1321 ACATCAACTTTTGGATTCAGAGTGGAGTGGCTGAGAGCAGTCTTGAATGATATCA 1380  
DB 1348 ACATCAACTTTTGGATTCAGAGTGGAGTGGCTGAGAGCAGTCTTGAATGATATCA 1407  
QY 1381 AGTATTTCTTCTTCATCACTCCCAAGAGACACATGACTGTGCT--ATGATCAAAAGC 1437  
DB 1408 AGTATTTCTTCTTCATCACTCCCAAGAGACACATGACTGTGCT--ATGATCAAAAGC 1467  
QY 1438 AGATGAATG--TTGCTGTGCTG--TTTGGGCTGTGTTCTTATGT--TGTGACAGATG 1494  
DB 1468 AGATGAATGTTGTGCTGTGCTGTTTGGGCTGTGTTCTTATGTGTTGACAGATG 1527  
QY 1495 AAGAAATGCTGCTGAGTCTTGAAGAAACGTAAGAAAGAA--CGTTTTCATGCTTGTGCC 1553  
DB 1528 AAGAAATGCTGCTGAGTCTTGAAGAAACGTAAGAAAGAA--CGTTTTCATGCTTGTGCC 1587  
QY 1554 --AGGAATCTGGGCTGCAACTTT--GAAAAACCTCTTCATCAATCAATTT--CATCC 1608  
DB 1588 CAGGAATCTGGGCTGCAACTTTGGAAGAAACCTCTTCATCAATCAATTT--CATCC 1647  
QY 1609 AATTCATCTTCAAGCAAACTCT--ATTCAATGCTTCT--GTATTAATCTTCTT--GATA 1665  
DB 1648 AATTCATCTTCAAGCAAACTCTTAATTTGATGCTTCTGTTATATCTTCTTGGATA 1707  
QY 1666 CTTTCCAAATCTCTG--ATTCTAAGAAAGAAATCAATTTCTCCCTCC--CTCCACCC 1719  
DB 1708 CTTTCCAAATCTCTGATTTCTAAGAAAGAAAGGAATCAATTTCTCCCTCCCTCCACCC 1767  
QY 1720 ACATAGATCAACATATGATGAGGATTAAGTGGGGGCA--TTTCTTATATCACTCTTA 1778  
DB 1768 ACATAGATCAACATATGATGAGGATTAAGTGGGGGCAATTTTCTTATATCACTCTTA 1827  
QY 1779 AAAACATTTGTTCCACTTTAAA--GTAAACATTTAAATTTTGAAGATCTGTAA 1836  
DB 1828 AAAACATTTGTTCCACTTTAAAAGTTAAACATTTAATTTTGAAGATTAATCTGAA 1887

Oy 1837 AAAAAAA 1844  
| | | | |  
Db 1886 AAAAAAA 1895

Search completed: December 22, 2003, 15:09:56  
Job time : 455.127 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:22:34 ; Search time 3623 Seconds  
(without alignments)  
12417.225 Million cell updates/sec

Title: US-09-745-763-35

Perfect score: 1851  
Sequence: 1 GGCTAGGCGCCGAGCTTAGT.....CTGAAAAAAAAAAAAAAAAAAAA 1851

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

```

EST:
1:  em_estba:**
2:  em_esthum:**
3:  em_estlin:**
4:  em_estmu:**
5:  em_estcov:**
6:  em_estpl:**
7:  em_estro:**
8:  em_hlc:**
9:  gb_est1:**
10: gb_est2:**
11:  gb_hlc:**
12:  gb_est3:**
13:  gb_est4:**
14:  gb_est5:**
15:  em_estfun:**
16:  em_estom:**
17:  em_gss_hum:**
18:  em_gss_inv:**
19:  em_gss_pln:**
20:  em_gss_vrt:**
21:  em_gss_fun:**
22:  em_gss_mam:**
23:  em_gss_mus:**
24:  em_gss_pro:**
25:  em_gss_rtd:**
26:  em_gss_phg:**
27:  em_gss_vrl:**
28:  gb_gss1:**
29:  gb_gss2:**

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1164.8	62.9	11	AK032972 Mus muscu
2	1164	62.9	11	BC010977 Mus muscu
3	1162.2	62.8	11	AK075686 Mus muscu
4	993.8	53.7	11	BC012019 Homo sapi

C	5	992.8	53.6	996	13	BX355939	BX355939
	6	985.8	53.3	1050	13	BX439467	BX439467
	7	940.8	50.8	1038	13	BX355940	BX355940
C	8	925.6	50.0	1020	13	BX439466	BX439466
	9	894	48.3	1073	13	BX360506	BX360506
C	10	872.4	47.1	1126	13	BX360507	BX360507
	11	847.8	45.7	909	13	BQ878966	BQ878966
C	12	845.2	45.8	1003	13	BX416895	BX416895
	13	840.8	45.4	1012	13	BX416896	BX416896
	14	838.6	45.3	919	13	BX335995	BX335995
C	15	831.6	44.9	929	13	BX335994	BX335994
	16	831	44.9	998	13	BX460463	BX460463
	17	772.6	41.7	945	12	B1909780	B1909780
	18	761.6	41.1	801	12	B1754143	603025609
	19	753.8	40.7	1081	13	BQ072892	AGENCOURT
	20	748.6	40.4	866	13	B1146905	AGENCOURT
	21	745.4	40.3	922	12	BG761741	602771936
	22	740	40.0	902	10	BG751497	602730076
	23	739.8	40.0	783	12	B1754468	603022607
	24	724.2	39.1	879	13	BQ218838	AGENCOURT
C	25	723.2	39.1	752	14	CD370462	CD370462
	26	719.8	38.9	832	12	B1911772	603065377
	27	719.2	38.9	739	14	CB958693	AGENCOURT
	28	716.4	38.7	786	10	BG430966	602500267
	29	709	38.3	944	12	B1819273	603037713
	30	691.4	37.4	774	14	CB308109	AGENCOURT
	31	689.8	37.3	1201	13	BX379674	BX379674
	32	688.8	37.2	806	12	B1764052	603043258
C	33	686.8	37.1	703	13	B0625199	UI-H-FG1-
	34	684	37.0	805	10	BG563740	602584524
	35	683.6	36.9	813	14	CD519131	AGENCOURT
C	36	681.8	36.8	708	14	CD366462	UI-H-FG1-
	37	676.4	36.5	691	14	CA436850	UI-H-DF1-
C	38	675.8	36.5	701	12	B1771109	603059666
	39	675.8	36.5	745	12	B1462884	603203005
	40	671.6	36.3	729	12	B1460469	603200966
	41	665.2	35.9	889	10	BE906771	601498329
	42	663.8	35.9	937	14	CD515740	AGENCOURT
	43	662.2	35.2	727	14	BG765433	602738911
	44	638.2	34.5	807	10	BG190926	RST10008
C	45	636.8	34.4	653	10	BE218907	hV46D06.X

## ALIGNMENTS

RESULT 1  
AK032972  
LOCUS  
DEFINITION  
Mus musculus 12 days embryo male wolffian duct includes surrounding region cDNA, RIKEN full-length enriched library, clone:6720483C06  
product:Plasma glutamate carboxypeptidase, full insert sequence.  
ACCESSION  
AK032972  
VERSION  
AK032972.1 GI:26328732  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
REFERENCE  
1  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)  
JOURNAL  
MEDLINE  
PUBMED  
10349636  
AUTHORS  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)  
JOURNAL  
MEDLINE  
PUBMED  
11042159



REFERENCE	AUTHORS
3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubaki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
REFERENCE	11076861
AUTHORS	4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nakai, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staudli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzei, J., Nombela, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohlschki, S. and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
REFERENCE	11217851
AUTHORS	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 1739)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sato, H., Sasaki, D., Shibata, K., Shinagawa, A., Shitaki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, Y., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
FEATURES	location/Qualifiers
	URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/

source	CDS
1..1739	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="FANTOM_DB:6720483C06"
	/db_xref="taxon:10090"
	/clone="6720483C06"
	/sex="male"
	/tissue_type="wolfian duct includes surrounding region"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="12 days embryo"
81..1493	/note="unnamed protein product; plasma glutamate carboxypeptidase (MCD) [MGI:1889205, GB NM_018755, evidence: B1875N_99% match=543]"
	putative"
	/codon_start=1
	/protein_id="BAC28105.1"
	/db_xref="GI:26328733"
	/translation="MRSLPFLVHLALGSGKAVPKNGVORTPREIKELANYEDV AKALINLAVVGKQNRSEYRLGLVDYVGRLLSGSKRLKALIMQNLQODGLENYH LEQVRIYHMERGESAMLEPRIRHRAILDLGSSIGTPPGITRAVLVVASPELQAR ASEARGLIVYNDPVTGYETQVIRVQGAVALKVGAVSLIQSVASFYSPTGIGI KYDGVKIPVPLACTIVEDAMSRMSRGRKIVIHLEMGAKTVPDTSFNTVAITGS MYDEEVVLVGLHSDMDVQGLDDGGAISWALSDVQLGRPRTLRILVLTAE EOGSIGASOYEELHAKNIISKYSLVLEADSGPLPTGLQPTGSKARAKMEKNNLQIP LNTKPSNKGCTDINPVIQAGVPGASLRDILKYFFPHSHSDITVMDPKQNNVA AYAVALVAVVADMDENLPRS"
	1722..1737
	/note="putative"
	1739
	/note="putative"
	polya_signal
	polya_site
BASE COUNT	482 a 369 c 423 g 465 t
ORIGIN	
Query Match	62.9%; Score 1164.8; DB 11; Length 1739;
Best Local Similarity	83.8%; Pred. No. 3.1e-211;
Matches 1329; Conservative 0; Mismatches 254; Indels 3; Gaps 1;	
QY	131 TGTTCACCTTTTATCCTGTGCTGTGGAAAAGCTATATGCAAGATGGCATCTTAAGAG 190
DB	107 TGTTCACCTTTTATCCTGTGCTGTGGAAAAGCTATATGCAAGATGGCATCTTAAGAG 166
QY	191 GACTTTGAAGAATAAAGAATAATAGCAGCTGTGGAGATGTTGCTTAAGAACATCAT 250
DB	167 AACATTGCGAATATTAAGAATAATAGCAGCTATGAAGATGTTCTTAAGAACATCAT 226
QY	251 CAACCTAGCTGTTTATGTTAAGCCAGAAACAATCCTATAGAGATGGCACTTCTGCT 310
DB	227 CAACCTTGTGTTTATGTTAATCCAGAACCGGCTCTATAGAGATGGCACTTCTGCT 266
QY	311 TGATACCTTTGAGCCAGACGTGAGTCTCCAGAAAGCTTGAAGAAAGCCATCCAAATAT 370
DB	287 TGATACCTTTGAGCCAGACGTGAGTCTTAAAGAACTTGAAGAAAGCTTTCAAATCAT 346
QY	371 GTACCAAAACCTGAGAGAGATGGGCTGAGAAAGTTCACCTGAGGACAGTGAATACC 430
DB	347 GTACCAAAACCTGAGAGAGATGGGCTGAGAAATGTTTCACTGAGGACAGTGAATACC 406
QY	431 CCACTGGAGAGGGGAGAGAAATCAGCTGTGATGCTGGAGCCAGAAATTCATTAAGATAGC 490
DB	407 CCACTGGAGAGGGGAGAGAAATCAGTGAATGAGTGTGAGCTTGACCTCGAATTCACAAAGATGGC 466
QY	491 CATCTGGGCTTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCATTTACAGAGAAATGCTT 550
DB	467 TATTTCTGTGCTTTGGCAGCAGCATTTGGGACTCTCCAGAGGATATCACAGAAAGTCT 526
QY	551 GGTGGTGAACCTTTGATGATGACAGAGAGAGGCTCTGAGAGCAAGAGAGAGATGTG 610
DB	527 GGTGGTGAACCTTTGATGATGACAGAGAGAGATGACAGAGAGAGAGATGATCAT 586
QY	611 TGTATTAAACAACCTTCACTCAACTCAAGAGAGGTCGAATACCGAAGCAGAGGAGGC 670

```

Db      587 TGTATTAAACGAGCTTACACTGCTATGAGAGACTGTGCACTACCGGGGTGCAGGGAGC 646
Qy      671 GGTGAAGCTGCGAAGGTGGGGCTTTGGCATCTCTCATTTGGATCCGGGCTCTCTTC 730
Db      647 TGTGAAGCTGCGAAGGTGGGGCTTTGGCATCTCTCATTTGGATCCGGGCTCTCTTC 706
Qy      731 CATCTACAGTCTGACACAGGTATTCAGAAATACAGAGTGGCTGCCAAATTCAC 790
Db      707 CATCTACAGTCTGACACAGGTATTCAGAAATACAGAGTGGCTGCCAAATTCAC 766
Qy      791 AGCTGTATTAAGGTGAGAGTGCAGAAATATGTCAGAAATGGCTTCTCATGGAGTCA 850
Db      767 AGCTGTATTAAGGTGAGAGTGCAGAAATATGTCAGAAATGGCTTCTCATGGAGTCA 826
Qy      851 AATTGTATTAAGGTGAGAGTGCAGAAATATGTCAGAAATGGCTTCTCATGGAGTCA 910
Db      827 AATTGTATTAAGGTGAGAGTGCAGAAATATGTCAGAAATGGCTTCTCATGGAGTCA 886
Qy      911 TGTAGCAGAGTCACTGGAGCAAAATATCCAGAACAGTTGTACTGTCACTGAGACATCT 970
Db      887 TGTAGCAGAGTCACTGGAGCAAAATATCCAGAACAGTTGTACTGTCACTGAGACATCT 946
Qy      971 GGAAGAGTGGAGTGTGGGCGAGGTGCATGATGATGCGGTGAGCTTTATATCATG 1030
Db      947 GGAAGAGTGGAGTGTGGGCGAGGTGCATGATGATGCGGTGAGCTTTATATCATG 1006
Qy      1031 GGAAGAGTGGAGTGTGGGCGAGGTGCATGATGATGCGGTGAGCTTTATATCATG 1090
Db      1007 GGAAGAGTGGAGTGTGGGCGAGGTGCATGATGATGCGGTGAGCTTTATATCATG 1066
Qy      1091 GCTCTGAGCTGACAGAGAACAGGTGAGTGTGCTTCCAGATTTATCAGTATACAA 1150
Db      1067 GCTCTGAGCTGACAGAGAACAGGTGAGTGTGCTTCCAGATTTATCAGTATACAA 1126
Qy      1151 GGTAAATATTTTCCAACTACAGTCTGTGTGATGATGATGATGATGATGATGATGAT 1210
Db      1127 GGTAAATATTTTCCAACTACAGTCTGTGTGATGATGATGATGATGATGATGATGAT 1186
Qy      1211 TGGGCTGCAATTCATCTGAGTGAAGAGGCGAGGCGCATGATGAGAGGTTATGAGCT 1270
Db      1187 TGGGCTGCAATTCATCTGAGTGAAGAGGCGAGGCGCATGATGAGAGGTTATGAGCT 1246
Qy      1271 GCTGAGAGCTCTCAATATCACTCAAGTCTGAGCATGAGAGGAGCAGACATCACTT 1330
Db      1247 TCTGCAACCTCTCAATATCACTCAAGTCTGAGCATGAGAGGAGCAGACATCACTT 1306
Qy      1331 TTGATCCAGAGTGGAGTGTGAGGCGAGTCTGATGATGATGATGATGATGATGATGAT 1390
Db      1307 CTGATCCAGAGTGGAGTGTGAGGCGAGTCTGATGATGATGATGATGATGATGATGAT 1366
Qy      1391 CTTTCATCACTCCACGAGACACCATGATGATGATGATGATGATGATGATGATGATGAT 1450
Db      1367 TTTTCATCACTCCACGAGACACCATGATGATGATGATGATGATGATGATGATGATGAT 1426
Qy      1451 TGTGCTGTGTGGAGTGTGTTTCTTATGTTGTCAGACATGAGAAATGTCGCTAG 1510
Db      1427 TGTGCTGTGTGGAGTGTGTTTCTTATGTTGTCAGACATGAGAAATGTCGCTAG 1486
Qy      1511 GTCTAGAAACAGTAAAGAAAGCTTTTCTGCTTGTGCGCAGAAATCTGCGTCTGC 1570
Db      1487 GTCTAGAAACAGTAAAGAAAGCTTTTCTGCTTGTGCGCAGAAATCTGCGTCTGC 1546
Qy      1571 AACTTTGAAAACCTCTTTCATACATTAATTTATCTCAATTTATCTTCAAGCAACT 1630
Db      1547 ATCTTTACACATCTCTTTCATACATTAATTTATCTCAATTTATCTTCAAGCAACT 1606
Qy      1631 CTATTTCACTGCTTCTGTTATTTATCTTCTGATCTTCTGATCTTCTGATCTTCTGAT 1690
Db      1607 TTTCTTATACCTTCTGTTATTTATCTTCTGATCTTCTGATCTTCTGATCTTCTGAT 1663
Qy      1691 AAGAATCATTTCTCCCTCCCTCC 1716
Db      1664 TAAATCATGATCCCTACAGCACACC 1689

```

```

RESULT 2
EC010977
LOCUS
DEFINITION
Mus musculus, similar to plasma glutamate carboxypeptidase, clone
IMAGE:4018296, mRNA.
ACCESSION
EC010977
VERSION
EC010977.1
KEYWORDS
HTC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1868)
REFERENCE
Strauberg, R.
Direct Submission
Submitted (23-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
NCBI-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcaps-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
http://www.systemsbio.org
contact: amadan@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketterman, Anuradha
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAC Plate: 18 Row: 9 Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 9055233
This clone has the following problem: retained intron.
FEATURES
Location/Qualifiers
source
1..1868
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4018296"
/tissue_type="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clone) outgrowth infected with the virus MMTV."
/clone_1kb="NCI CGAP_Lu29"
/lab_host="DH10B"
/label="Vector: PCMV-Spore6"
BASE COUNT
505 a 418 c 460 g 485 t
ORIGIN
Query Match 62.9%; Score 1164; DB 11; Length 1868;
Best Local Similarity 83.5%; Pred. No. 4.3e-211;
Matches 1333; Conservative 0; Mismatches 260; Indels 3; Gaps 1;
121 TTTTGGTGTGTTACCTTTTATCCCTGTGCTGAGAAAGTATGCAAGATGGCA 180
217 TTCTGTATCATGTTACCTTTTACCTTGGCTGTGTAAGCTGTATTCAGAAATGTG 276
181 TCTTGAAGAGACTTTTGAAGAAATGAAGAAATAGCCAGTGTGAGATGTGCTA 240
277 TTCTCAGCAACATTTTCGAAGAAATGAAGAAATAGCCAACTATGAAGATGTCTA 336
241 AAGCAATCATCACTGCTTTATGTTATGTTAAAGCCGAGAACATCTTATGAGGATGG 300
337 AAGCAATTCACCTGTTGTTATGTTATGTTAAATACGAAACGGTCTTATGAGGATGG 396
301 CACTTGTGTTACTGTTGAGACCCAGACTGAGTGTCTCAAGAACCTGAGAAAGCA 360

```

Db 397 GACTTCTAGATTGATCTGTGGAGCCAGACTGAGTGGCTCTAAGAACCTAGAGAAAGCTA 456  
 Qy 361 TCCAAATTAATGACAAAACCTGACGAAAGAGGCTGGAGAAATTCACCTGGAGCCAG 420  
 Db 457 TCCAAATGATGACAAAACCTGACGAAAGAGGCTGGAGAAATTCACCTGGAGCCAG 516  
 Qy 421 TGAGAAATACCCACCTGGAGAGAGAGAAATTCAGCTGTGTGAGTGGAGCCAAATTC 480  
 Db 517 TCAGAAATACCCACCTGGAGAGAGAGAAATTCAGCTGTGTGAGTGGAGCCAAATTC 576  
 Qy 481 ATAGATAGCCATCTGGAGTCTGGAGAGAGAGAAATTCAGCTGTGTGAGTGGAGCCAAATTC 540  
 Db 577 ACAGATAGCCATCTGGAGTCTGGAGAGAGAGAAATTCAGCTGTGTGAGTGGAGCCAAATTC 636  
 Qy 541 CAGAGTCTGGAGTCTGGAGTCTGGAGAGAGAGAAATTCAGCTGTGTGAGTGGAGCCAAATTC 600  
 Db 637 CAGAGTCTGGAGTCTGGAGTCTGGAGAGAGAGAAATTCAGCTGTGTGAGTGGAGCCAAATTC 696  
 Qy 601 GGAAGATTTGTTTATTAACCAACCTTACATCACTCAAGAGAGAGAGAGAGAGAGAGAG 660  
 Db 697 GGAAGATTTGTTTATTAACCAACCTTACATCACTCAAGAGAGAGAGAGAGAGAGAGAG 756  
 Qy 661 GCGAG 720  
 Db 757 TCGAG 816  
 Qy 721 CCT 780  
 Db 817 CTTCT 876  
 Qy 781 AAATTCACCAAG 840  
 Db 877 AGATTCACCAAG 936  
 Qy 841 ATGGATGCAAAATTTGTCATTCAGCTTAAAGATGGGGGCAAAAGCTTACCCAGATCTGATT 900  
 Db 937 GTGGGCAAAATTTGTCATTCAGCTTAAAGATGGGGGCAAAAGCTTACCCAGATCTGATT 996  
 Qy 901 CCTTAAACAGCTGTACAGAGATCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
 Db 997 CCTTAAACAGCTGT 1056  
 Qy 961 GTGACATCTGGAG 1020  
 Db 1057 GTGACATCTGGAG 1116  
 Qy 1021 TTATATCATGGAG 1080  
 Db 1117 TCATATCATGGAG 1176  
 Qy 1081 TGGGCTGT 1140  
 Db 1177 TGGGCTGT 1236  
 Qy 1141 AGTTACACAG 1200  
 Db 1237 AGTTACACAG 1296  
 Qy 1201 TCTTACACAG 1260  
 Db 1297 TCTTACACAG 1356  
 Qy 1261 TTATAGAGCTGT 1320  
 Db 1357 TCATATCATGGAG 1416  
 Qy 1321 ACATCAACTTTTGTATTCACAG 1380  
 Db 1417 ATATATTAATCTGT 1476  
 Qy 1381 AGTATTTCTTCTTCATCACTGT 1440  
 Db 1477 AGTATTTCTTCTTCTTCATCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1536

Qy 1441 TGAATTTGCTGT 1500  
 Db 1337 TGAATTTGCTGT 1556  
 Qy 1501 TGTGCTTGT 1560  
 Db 1597 TGTGCTTGT 1656  
 Qy 1561 CTGGGCTGT 1620  
 Db 1657 CCAACTGT 1716  
 Qy 1621 AAGCACTGT 1680  
 Db 1717 GGGCTGT 1773  
 Qy 1681 GATTCTAGAAAAAGAAATCATCTTCCCTCCCTCC 1716  
 Db 1774 GTTCTAGAAATTAATCATGATCCTACTGACACACC 1809  
 RESULT 3  
 AK075686  
 LOCUS  
 DEFINITION  
 Mus musculus 1805 bp mRNA linear HTC 07-DEC-2002  
 enriched library, clone:1190003p12 product:plasma glutamate  
 carboxypeptidase, full insert sequence.  
 ACCESSION  
 AK075686  
 AK075686.1 GI:26344480  
 VERSION  
 HTC; CAP trapper.  
 KEYWORDS  
 Mus musculus (house mouse)  
 SOURCE  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1  
 Carninci, P. and Hayashizaki, Y.  
 TITLE  
 High-efficiency full-length cDNA cloning  
 JOURNAL  
 Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE  
 99279253  
 PUBMED  
 10349636  
 REFERENCE  
 2  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL  
 Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE  
 20499374  
 PUBMED  
 11042159  
 REFERENCE  
 3  
 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,  
 Komno, H., Akiyama, T., Nishi, K., Kikunishi, T., Tashiro, H., Itoh, M.,  
 Aizawa, K., Iizawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishigaki, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Iwama, M., Ohara, E., Matsumoto, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 TITLE  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multiplexed sequencer  
 JOURNAL  
 Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE  
 20530913  
 PUBMED  
 11076861  
 REFERENCE  
 4  
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
 Aizawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S.,  
 Aizawa, K., Iizawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,  
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
 Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pebole, G.,  
 Quackenbush, J., Schiml, L.M., Staudt, F., Suzuki, R., Tomita, M.,  
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,

Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Feltcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hotam, M., Hume, D.A., Kamuya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Marzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Saeki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wyshaw-Borja, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.

**TITLE** Functional annotation of a full-length mouse cDNA collection

**JOURNAL** Nature 409 (6821), 685-690 (2001)

**MEDLINE** 21085660

**PUBMED** 11217851

**REFERENCE** 5

**AUTHORS** The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

**TITLE** Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

**JOURNAL** Nature 420, 563-573 (2002)

**REFERENCE** 6 (bases 1 to 1805)

**AUTHORS** Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arahawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komuro, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Tejima, Y., Toy, T., Yamamura, T., Yamana, I., Yasunaka, T., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

**TITLE** Direct Submission

**JOURNAL** Submitted (16-ARR-2002) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome-gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

**COMMENT** cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

**FEATURES** Please visit our web site for further details.

**SOURCE** URL: http://genome-gsc.riken.go.jp/

location/Qualifiers

1..1805

organism="Mus musculus"

molecule="mRNA"

strain="C57BL/6J"

db\_xref="PANTOM DB:1190003P12"

db\_xref="MGI:1909852"

db\_xref="taxon:10090"

clone="1190003P12"

tissue\_type="whole body"

clone\_id="RIKEN full-length enriched mouse cDNA library"

dev\_stage="18-day embryo"

147..1559

note="unmated protein product; plasma glutamate carboxypeptidase (MCP|MGI:1889205, GI|NM\_018755, evidence: BLASTN, 99%, match=943)

putative"

codon\_start=1

protein\_id="BAC35891.1"

db\_xref="MGI:2634481"

db\_xref="MGI:1923585"

translation="MRLSLFLFIYHLALSGSKAVFKNGVSQRFREIKETIANEDVAKAIINLAVGKYQNSRYERGLLVDTVPRLSGSKLEKAIQMYNLODGLNVHLEOVYPMERGESAVMLBPRIHKMLIGGSSITGPGGITLVEVAVASFDLQRRASEARGLIIVNPGYTGKTYQYRGAVBAKAVAAVSLIGVASFSIYSPHTGIG

KYDGVKPIPTACTIVEDAEWMSRMNRKNKIVIHMEKATYPTDSEFNTVAETGS  
MYEEFVLVSGHDSMDVGGALDDGGAFI SWEALSLVDLGRKRTLRVLMTAE  
BOEGEIOYVEYEHKANKISKYSLVMEADSGTFPLPTGLOFSGKARAIKEVNNLOP  
LNTKXPSNEGTDINFWIOAGVPGASISDDLYKFFPHSHSDMTVMDDPKOMNVA  
AVAAVAVYVYVADMDENLPRS"

polya\_signal  
1788..1793  
note="putative"

polya\_site  
1805  
note="putative"

BASE COUNT 486 a 397 c 445 g 476 t 1 others

ORIGIN

Query Match 62.8%; Score 1162.2; DB 11; Length 1805;  
Best Local Similarity 83.7%; Pred. No. 9.5e-211;  
Matches 1327; Conservative 0; Mismatches 256; Indels 3; Gaps 1;

QY 131 TGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGATGCATCTTAAG 190  
DB 173 TGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGATGCATCTTAAG 232  
QY 191 GACTTTGAAGAAATTAAGAAATAGCCAGCTGTGGAGATGTTGCTAAAGCAATCAT 250  
DB 233 AACATTCGAGAAATTAAGAAATAGCCAGCTATGAAGATGTTGCTAAAGCAATCAT 292  
QY 251 CAACCTAGCTGTTTATGTTAAAGCCAGAAAGATCCTATGAGCGATTCGCACTTCTGT 310  
DB 293 CAACCTAGCTGTTTATGTTAAATACCAAGACCGCTCTATGAGCGTTTGGACCTTCTAGT 352  
QY 311 TGATACCTGTTGAAGCCAGACTGATGCTGCTCAAGAACTTGAAGAAAGCAATTCAT 370  
DB 353 TGAATACCTGTTGAAGCCAGACTGATGCTGCTCAAGAACTTGAAGAAAGCAATTCAT 412  
QY 371 GTACCAAAACCTGAGCAAGATAGGCGTGGAGAAAGTTACCTGAGCCATGAGAAATAC 430  
DB 413 GTACCAAAACCTGAGCAAGATAGGCGTGGAGAAAGTTACCTGAGCCATGAGAAATAC 472  
QY 431 CCACCTGAGAGAGGAGAGAAATCAGCTGATGATGCTGAGCCCAAGAAATTCATTAAGATAC 490  
DB 473 CCACCTGAGAGAGGAGAGAAATCAGCTGATGATGCTGAGCCCTGAAATTCATTAAGATAC 532  
QY 491 CATCTGGGCTGTTGGAGAGCAATGGGACTCTCCGAAAGGCAATTCAGAGAAATGCT 550  
DB 533 TATCTGAGCTGTTGGAGAGCAATGGGACTCTCCGAAAGGCAATTCAGAGAAATGCT 592  
QY 551 GGTGCTGACCTCTTTCATGATGATGCAAGAGAGGCTTCAGAAAGCAAGAGAAATGCT 610  
DB 593 GGTGCTGACCTCTTTCATGATGATGCAAGAGAGGCTTCAGAAAGCAAGAGAAATGCT 652  
QY 611 TGTTCATACCAACCTTACATCACTCAAGAGAGGCTCAATACCAAGAGAGAGGAGG 670  
DB 653 TGTTCATACCAACCTTACATCACTCAAGAGAGGCTCAATACCAAGAGAGAGGAGG 712  
QY 671 GGTGAGAGCTGCAAGAGTGGGAGCTTGGGATCTTCATTCATTCATTCATTCATTC 730  
DB 713 TGTGAGAGCTGCAAGAGTGGGAGCTTGGGATCTTCATTCATTCATTCATTCATTC 772  
QY 731 CATTCATGCTTCATCAACAGATATTCAGAAATACAGAGAGGCTGCTCCAAATTCAC 790  
DB 773 CATTCATGCTTCATCAACAGATATTCAGAAATACAGAGAGGCTGCTCCAAATTCAC 832  
QY 791 AGCTGATATACAGGATGAGATGAGAAATGATGCAAGATGAGCTTCATGAGGATCAA 850  
DB 833 AGCTGATATACAGGATGAGATGAGAAATGATGCAAGATGAGCTTCATGAGGATCAA 892  
QY 851 AATGTCATTCAGCTAAAGATGGGAGCAAGACCTATCCAGATATTCATTCATTCAT 910  
DB 893 AATGTCATTCAGCTAAAGATGGGAGCAAGACCTATCCAGATATTCATTCATTCAT 952  
QY 911 TGTGAGAGATTCATTCAGAGCAATATTCAGAAAGAGGCTTGTACTGCTAGTGAACAT 970  
DB 953 TGTGAGAGATTCATTCAGAGCAATATTCAGAAAGAGGCTTGTACTGCTAGTGAACAT 1012  
QY 971 GGAAGAGTGGAGATGTTGGAGAGGCTGCATGATGATGAGGCTTATATCATG 1030

```

Db      1013  GGACAGCTGGAGTGTGGGCGAGGTCACATGATGATGCTGGAGCCTTCAATATCATG 1072
Qy      1031  GGAAGACCTCTCACTATTAAAGATCTGGGCTGGCTCCAAAGAGACCTCGGCTGCT 1090
Db      1073  GGAAGACCTCTCACTATTAAAGATCTGGGCTGGCTCCAAAGAGACCTCGGCTGCT 1132
Qy      1091  GCTCTGACATGACAGAAACAAGAGTGGATGGTCTCCAGTATTATCACTACACA 1150
Db      1133  GCTCTGACATGACAGAAACAAGAGTGGATGGTCTCCAGTATTATCACTACACA 1192
Qy      1151  GGTAAATATTTCCAACTACAGTCTGGTGAAGTCTGACGAGAACCTTCTTACCC 1210
Db      1193  GGCAATATTTCCAACTACAGTCTGGTGAAGTCTGACGAGAACCTTCTTACCC 1252
Qy      1211  TGGGCTGCAATTCACCTGAGAGTGAAGAAAGGCGGCGCATCATGAGAGCTATAG 1270
Db      1253  TGGAGCTGAGTTCACCTGAGAGTGAAGAAAGGCGGCGCATCATGAGAGCTATAG 1312
Qy      1271  GCTGAGAGCCCTCAATATCACTCAGGCTCTGAGCCATGAGAGGAGACATCACT 1330
Db      1313  TCTGCAACCCCTCAATATCACTCAGGCTCTGAGCCATGAGAGGAGACATCACT 1372
Qy      1331  TTGATATCAAGCTGAGAGTCTGAGCCATGAGAGGAGACATCACTATTTCTT 1390
Db      1373  CTGATATCAAGCTGAGAGTCTGAGCCATGAGAGGAGACATCACTATTTCTT 1432
Qy      1391  CTTCATCATCTCCACGAGAGACATCATGATCTGATGATCCAAAGCAGATATGCT 1450
Db      1433  TTTTCATCATCTCCACGAGAGACATCATGATCTGATGATCCAAAGCAGATATG 1492
Qy      1451  TGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGACATGAGAAATGCTGCTAG 1510
Db      1493  TGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGACATGAGAAATGCTGCTAG 1552
Qy      1511  GTCTGAGAAACAGTAAAGAAAGAGTTTCATGCTTCTGCGCAGAGATCTCGGCT 1570
Db      1553  GTCTGAGAGAAACAGTAAAGAGAGAGACCTTCTCTGTAAGTGGAAATCCCACT 1612
Qy      1571  AACTTGGAGAAACCTCTCTTCACTATCAATTCATTCATCTTCAAGCAACT 1630
Db      1613  AACTTGGAGAAACCTCTCTTCACTATCAATTCATTCATCTTCAAGCAACT 1672
Qy      1631  CAATTCATCAAGCTTCTGTTATATCTTCTGATCACTTCCAAATCTCTGATCTAG 1690
Db      1673  TTCTTATACCTTCTGTTATATCTTCTGATCACTTCTTCAATCTTCTTCAAG 1729
Qy      1691  AAAAGATCATTTCTCCCTCCCTCC 1716
Db      1730  TAAATCATGATCCCTACTGCAACACC 1755

RESULT 4
BC012019      1024 bp      mRNA      linear      HTC 04-MAR-2003
LOCUS      Homo sapiens, Similar to plasma glutamate carboxypeptidase, clone
DEFINITION      IMAGE:4455631, mRNA.
ACCESSION      BC012019
VERSION      BC012019.1
KEYWORDS      GI:15080558
SOURCE      HTC.
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 1024)
AUTHORS      Strausberg, R.
TITLE      Direct Submission
JOURNAL      Submitted (30-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK      NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT      Contact: MGC help desk

```

```

Email: cga@bcm.tmc.edu
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.bhsc.bcm.tmc.edu/cdna/
Contact: am@bcm.tmc.edu
Gunnar, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H.,
Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.W., Nantavali,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://lml.gov
Series: IRAC Plate: 28 Row: b Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5174626
This clone has the following problem: retained intron.

FEATURES
    source
        1..1024
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4455631"
            /tissue_type="kidney, hypernephroma"
            /clone_id="NIH MGC_89"
            /lab_host="DH10B"
            /note="Vector: pCMV-Sport6"

BASE COUNT      317 a      210 c      256 g      241 t
ORIGIN
Query Match      53.7%; Score 993.8; DB 11; Length 1024;
Best Local Similarity 99.8%; Pred. No. 1.1e-178;
Matches 995; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

64 GATTATCTTAAACAAGAAACCACTGGAAGAAATGAATTCCTTATCTTCCATTTT 123
1 GATTATCTTAAACAAGAAACCACTGGAAGAAATGAATTCCTTATCTTCCATTTT 60
Qy      124  TCGGTGTTACCTTTTATCCCTGCTCTGCTGGAAGAGTATATGCAAGATGCAATCT 183
Db      61  TCGGTGTTACCTTTTATCCCTGCTCTGCTGGAAGAGTATATGCAAGATGCAATCT 120
Qy      184  CTAAGAGAGCTTTGAAGAAATTAAGAAATAGCAGCTGAGATGCTGTAAG 243
Db      121  CTAAGAGAGCTTTGAAGAAATTAAGAAATAGCAGCTGAGATGCTGTAAG 180
Qy      244  CAATCATCAACTGACTGTTATGTTAAGCCCAAGACAGATCTTATGAGCATTTGGCAC 303
Db      181  CAATCATCAACTGACTGTTATGTTAAGCCCAAGACAGATCTTATGAGCATTTGGCAC 240
Qy      304  TTCTGTTGATGATCTGTTGAGCCAGACTGAGTGGCTCCAAAGACTTGAAGAACCATTC 363
Db      241  TTCTGTTGATGATCTGTTGAGCCAGACTGAGTGGCTCCAAAGACTTGAAGAACCATTC 300
Qy      364  AAATTAATGACCAAAACCTGAGCAAGATAGGCTGGAAGAAAGTTCACTGAGCACTGA 423
Db      301  AAATTAATGACCAAAACCTGAGCAAGATAGGCTGGAAGAAAGTTCACTGAGCACTGA 360
Qy      424  GAATATCCCACTGGAGAGAGGAGAAAGATCAGCTGATGCTGAGACCAAGATTTGATA 483
Db      361  GAATATCCCACTGGAGAGAGGAGAAAGATCAGCTGATGCTGAGACCAAGATTTGATA 420
Qy      484  AGATAGCCATCTGGGCTTTGAGCAGCAATTTGGGCTCTCCAGAGGATTTACAGCAG 543
Db      421  AGATAGCCATCTGGGCTTTGAGCAGCAATTTGGGCTCTCCAGAGGATTTACAGCAG 480
Qy      544  AAGTTCTGAGTGAACCTTTTGAATGATGATGATGATGATGATGATGATGATGATG 603
Db      481  AAGTTCTGAGTGAACCTTTTGAATGATGATGATGATGATGATGATGATGATGATG 540
Qy      604  AGATTTGTTTATTAACAACCTTATCACTAAGAGAGGCTGCAATACCGAAGCGC 663

```

```

Db      541 AGATTGTTTATATACCACTTACATCACTACTACAGGACGGTGCATATCCGAACGC 600
Qy      664 AGGGGGGGGTGAGAGCTGCCAAGTGGGGGCTTTGGCATCTCTCATTCGATCCGTGGCCT 723
Db      601 AGGGGGGGGTGAGAGCTGCCAAGTGGGGGCTTTGGCATCTCTCATTCGATCCGTGGCCT 660
Qy      724 CCTTCTCATCTACAGTCTCCACACAGGTATTCAGAAATACAGAGATGGCGTCCCAAAA 783
Db      661 CCTTCTCATCTACAGTCTCCACACAGGTATTCAGAAATACAGAGATGGCGTCCCAAGA 720
Qy      784 TTCCAAGAGCTGTATTACGGTGGAGAGTGCAGAAATGATGTCAAGATGGCTTTCATG 843
Db      721 TTCCAAGAGCTGTATTACGGTGGAGAGTGCAGAAATGATGTCAAGATGGCTTTCATG 780
Qy      844 GGATCAAAATTTGATTCAGTAAAGATGGGGGCAAAAGCTTACCCAGATCTGATTCCT 903
Db      781 GGATCAAAATTTGATTCAGTAAAGATGGGGGCAAAAGCTTACCCAGATCTGATTCCT 840
Qy      904 TCAACACTGTAGACAGATCACTGGAGCAATATCCAGACAGTTGTACTGGTCAAGT 963
Db      841 TCAACACTGTAGACAGATCACTGGAGCAATATCCAGACAGTTGTACTGGTCAAGT 900
Qy      964 GACATCTGACAGCTGGAGTGTGGGCAAGGTGCCATGATGATGGCGTGGAGCCTTTA 1023
Db      901 GACATCTGACAGCTGGAGTGTGGGCAAGGTGCCATGATGATGGCGTGGAGCCTTTA 960
Qy      1024 TATCATGGGAAGCACTCTCACTTATTAAGATCTTTGG 1060
Db      961 TATCATGGGAAGCACTCTCACTTATTAAGATCTTTGG 997

```

RESULT 5  
BX355939/c 996 bp mRNA linear EST 05-MAY-2003

LOCUS BX355939 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA  
DEFINITION clone CSODI006Y115 3-PRIME, mRNA sequence.

ACCESSION BX355939  
VERSION BX355939.1 GI:30382027

KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 996)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished

COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4663.f For  
more information about this cluster, see

http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CSODI006AB08NP1&cluster=4663.f. Contact :  
Peng Liang Email : lliang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CSODI006AB08NP1.  
Location/Qualifiers

FEATURES  
SOURCE 1..996

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODI006Y115"

/cuisine\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NciI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 270 a 231 c 223 g 269 t 3 others

```

ORIGIN
Query Match 53.64; Score 992.8; DB 13; Length 996;
Best Local Similarity 99.64; Pred. No. 1.7e-178;
Matches 992; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy      788 AACAGCCTGTATTACGGTGGAGAGTGCAGAAATGATGTCAAGATGGCTTCTCATGGAT 847
Db      996 AACAGCCTGTATTACGGTGGAGAGTGCAGAAATGATGTCAAGATGGCTTCTCATGGAT 937
Qy      848 CAAATTTGATTCAGCTTAAAGATGGGGGCAAAAGACTTACCAGATCTGATTCCTTCAA 907
Db      936 CAAATTTGATTCAGCTTAAAGATGGGGGCAAAAGACTTACCAGATCTGATTCCTTCAA 877
Qy      908 CACTGTAGCAGAGTCACTGGGAGCAAAATTCAGAAACAGTTGTACTGTGAGTGCAGCA 967
Db      876 CACTGTAGCAGAGTCACTGGGAGCAAAATTCAGAAACAGTTGTACTGTGAGTGCAGCA 817
Qy      968 TCTGACAGCTGGAGATTTGGGAGAGGTGCCATGATGATGGCGTGGAGCCTTTATATC 1027
Db      816 TCTGACAGCTGGAGATTTGGGAGAGGTGCCATGATGATGGCGTGGAGCCTTTATATC 757
Qy      1028 ATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTTGGCGCT 1087
Db      756 ATGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTTGGCGCT 697
Qy      1088 GGTGCTCTGACAGTGCAGAAAGCAAGGTGAGTTGGTCCCTTCAGATTAATACATTACA 1147
Db      696 GGTGCTCTGACAGTGCAGAAAGCAAGGTGAGTTGGTCCCTTCAGATTAATACATTACA 637
Qy      1148 CAAGGTAAATATTTCACACTACAGTCTGTGATGATGATGATGATGATGATGATGATGAT 1207
Db      636 CAAGGTAAATATTTCACACTACAGTCTGTGATGATGATGATGATGATGATGATGATGAT 577
Qy      1208 CACTGGCTGCAATTCATCTGCGAGTGAAGAAAGCCAGGGCCATCATGAGAGAGTTATGAG 1267
Db      576 CACTGGCTGCAATTCATCTGCGAGTGAAGAAAGCCAGGGCCATCATGAGAGAGTTATGAG 517
Qy      1268 CTTGCTGACAGCCCTCAATATCACTGAGTCTGAGGCAATGAGAAAGGACAGCATCAA 1327
Db      516 CTTGCTGACAGCCCTCAATATCACTGAGTCTGAGGCAATGAGAAAGGACAGCATCAA 457
Qy      1328 CTTTGGATCCAAAGCTGAGTGGCTGGAGCAGTCTCTTATGATGATGATGATGATGATGAT 1387
Db      456 CTTTGGATCCAAAGCTGAGTGGCTGGAGCAGTCTCTTATGATGATGATGATGATGATGAT 397
Qy      1388 CTTTCTCATCACTCCACAGGAGACACCATGATGATGATGATGATGATGATGATGATGAT 1447
Db      396 CTTTCTCATCACTCCACAGGAGACACCATGATGATGATGATGATGATGATGATGATGAT 337
Qy      1448 TGTCTGCTGCTGTTGGGCTGTTGTTTCTTATGTTGTTGTCAGACATGGAAGAAATGCTGCC 1507
Db      336 TGTCTGCTGCTGTTGGGCTGTTGTTTCTTATGTTGTTGTCAGACATGGAAGAAATGCTGCC 277
Qy      1508 TAGTCTCTAGAAACAGTAGAAGAAAGCTTTTATGATGATGATGATGATGATGATGATGATG 1567
Db      276 TAGTCTCTAGAAACAGTAGAAGAAAGCTTTTATGATGATGATGATGATGATGATGATG 217
Qy      1568 TGCACATTTGAAAACCTCTCTTACATTAACATTTTATTCATTCATTCATTCATTCATTCAT 1627
Db      216 TGCACATTTGAAAACCTCTCTTACATTAACATTTTATTCATTCATTCATTCATTCATTCAT 157
Qy      1628 ACTCTATTTATGCTTTCTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1687
Db      156 ACTCTATTTATGCTTTCTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 97
Qy      1688 GAAAAGGAATCATTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 1747
Db      96 GAAAAGGAATCATTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 37
Qy      1748 CAGTGGGGGCAATTTCTTATATATCACTCTTAAAAAC 1783
Db      36 CAGTGGGGGCAATTTCTTATATATCACTCTTAAAAAC 1

```

RESULT 6  
 EX439467  
 LOCUS  
 DEFINITION BX439467 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE010YJ06  
 5-PRIME, mRNA sequence.  
 ACCESSION BX439467  
 VERSION BX439467.1 GI:30777746  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 1050)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 4663.f for  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DE010DE030P1cluster=4663.f. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0DE010DE030P1.  
 Location/Qualifiers  
 1..1050  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE010YJ06"  
 /issue\_type="PLACENTA"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with Not I and cloned into  
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."  
 BASE COUNT 307 a 233 c 271 g 239 t  
 ORIGIN  
 Query Match 53.3%; Score 985.8; DB 13; Length 1050;  
 Best Local Similarity 99.8%; Pred. No. 3,6e-177;  
 Matches 987; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Oy 19 GTCCGGAGAGCGCGCTCGCGCCGCGCAGAGCGCGCTTTCAGATTATCTTAACAG 78  
 Db 62 GTCCGGAGAGCGCGCTCGCGCCGCGCAGAGCGCGCTTTCAGATTATCTTAACAG 121  
 Oy 79 AAAACCACTGGAAAAAATGAATTCCTATCTTCGATTTTTCGTTGTTTACC 138  
 Db 122 AAAACCACTGGAAAAAATGAATTCCTATCTTCGATTTTTCGTTGTTTACC 181  
 Oy 139 TTTTATCCCTGCTCTGGGAAAGCTATATGCAAGATGCGATCTTAAAGAGACTTTTG 198  
 Db 182 TTTTATCCCTGCTCTGGGAAAGCTATATGCAAGATGCGATCTTAAAGAGACTTTTG 241  
 Oy 199 AAGAAATAAAGAAATAGCCACTGTGAGATGTTGCTTAAGCAATCATCAACCTAG 258  
 Db 242 AAGAAATAAAGAAATAGCCACTGTGAGATGTTGCTTAAGCAATCATCAACCTAG 301  
 Oy 259 CTGTTTATGTTAAAGCCGAGAACAGATCTTATGAGCGATTGSCACTTTCGTTGTTACTG 318  
 Db 302 CTGTTTATGTTAAAGCCGAGAACAGATCTTATGAGCGATTGSCACTTTCGTTGTTACTG 361  
 Oy 319 TTGACCCAGACTGAGTGGCTTCAAGAACCTTGAAGAAAGCCATCCAAATATATGACCAA 378  
 Db 362 TTGACCCAGACTGAGTGGCTTCAAGAACCTTGAAGAAAGCCATCCAAATATATGACCAA 421

Oy 379 ACTGCAGCAAGATGGGCTGGAGAAAGTTCACTTGAGCCAGTGAGATACCCCACTGGG 438  
 Db 422 ACTGCAGCAAGATGGGCTGGAGAAAGTTCACTTGAGCCAGTGAGATACCCCACTGGG 481  
 Oy 439 AGAGGGGAGAAAGATTCAGCTGTGTGAGCGCCAGAAATTCATAGATAGCCATCTGG 498  
 Db 482 AGAGGGGAGAAAGATTCAGCTGTGTGAGCGCCAGAAATTCATAGATAGCCATCTGG 541  
 Oy 499 GTCTTGGCAGAGCATTTGGGACTCTCTCCAGAAAGCATTTACAGCAAGTTCTGGTGTGA 558  
 Db 542 GTCTTGGCAGAGCATTTGGGACTCTCTCCAGAAAGCATTTACAGCAAGTTCTGGTGTGA 601  
 Oy 559 CCTCTTTGATGAATCTGAGAGAGGCGCTCAGAGCAAGAGGAGATTTGTTTATA 618  
 Db 602 CCTCTTTGATGAATCTGAGAGAGGCGCTCAGAGCAAGAGGAGATTTGTTTATA 661  
 Oy 619 ACCAACCTTACATCACTACTCAAGACGGTGCAATACGAAACGAGGGGCGGTGGAAG 678  
 Db 662 ACCAACCTTACATCACTACTCAAGACGGTGCAATACGAAACGAGGGGCGGTGGAAG 721  
 Oy 679 CTGCGAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGGCTCTTCTCCATCTACA 738  
 Db 722 CTGCGAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGGCTCTTCTCCATCTACA 781  
 Oy 739 GTCCCTCACAGCATTTACAGGAATACAGAGATGGCGGCCAAATTCGAACAGCTGTA 798  
 Db 782 GTCCCTCACAGCATTTACAGGAATACAGAGATGGCGGCCAAATTCGAACAGCTGTA 841  
 Oy 799 TTACGGTGAAGATGCAAGAAATGATGTCAGAAATGGCTTCTCATGGGATCAAAATTTGCA 858  
 Db 842 TTACGGTGAAGATGCAAGAAATGATGTCAGAAATGGCTTCTCATGGGATCAAAATTTGCA 901  
 Oy 859 TTACGCTTAAGATGGGGGCAAGACCTTACCAAGTATGATTTCTTGAACCTGTACAG 918  
 Db 902 TTACGCTTAAGATGGGGGCAAGACCTTACCAAGTATGATTTCTTGAACCTGTACAG 961  
 Oy 919 AGATCACTGGGAGCAAAATATCAGAACAGGTTGATGTCAGTGAAGCATCTGGACAGCT 978  
 Db 962 AGATCACTGGGAGCAAAATATCAGAACAGGTTGATGTCAGTGAAGCATCTGGACAGCT 1021  
 Oy 979 GGAGTGTGGGACAGGTCCTCATGATGAT 1007  
 Db 1022 GGAGTGTGGGACAGGTCCTCATGATGAT 1050  
 RESULT 7  
 EX355940  
 LOCUS  
 DEFINITION BX355940 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS0D1006Y115 5-PRIME, mRNA sequence.  
 ACCESSION BX355940  
 VERSION BX355940.1 GI:30384019  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 1038)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 4663.f for  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0D1006Y115cluster=4663.f. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600



FEATURES  
Location/Qualifiers  
1. 1038  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1006Y115"  
/issue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_id="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

Faraday Avenue Genoscope sequence ID : CS0D1006A080P1.  
BASE COUNT 304 a 237 c 262 g 233 t 2 others  
ORIGIN

Query Match 50.8%; Score 940.8; DB 13; Length 1038;  
Best Local Similarity 99.7%; Pred. No. 1.3e-168;  
Matches 953; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

2 GCTAGGCGGAGGCTTGTCTGAGAGCGGCTCGTCCGCGGCTAGAGCCGCTAT 61  
83 GCTAGGCGGAGGCTTGTCTGAGAGCGGCTCGTCCGCGGCTAGAGCCGCTAT 142  
62 CAGATTATCTTACAGAGAAACCACTGGAAGAAATGAATTCCTTATCTTGCAAT 121  
143 CAGATTATCTTACAGAGAAACCACTGGAAGAAATGAATTCCTTATCTTGCAAT 202  
122 TTTGGTGTGTTCACCTTTTATCCCTGTGCTCTGGAGAGCTTATGCAAGATGCGAT 181  
203 TTTGGTGTGTTCACCTTTTATCCCTGTGCTCTGGAGAGCTTATGCAAGATGCGAT 262  
182 CTCTAAGAGGACTTTTGAAGAAATGAAGAAATGCGAGCTGAGATGTTGCTAA 241  
263 CTCTAAGAGGACTTTTGAAGAAATGAAGAAATGCGAGCTGAGATGTTGCTAA 322  
242 AGCAATCATCAACCTAGCTGTTTATGTTAAGAGCCAGAACGATCTTATGAGCATGGC 301  
323 AGCAATCATCAACCTAGCTGTTTATGTTAAGAGCCAGAACGATCTTATGAGCATGGC 382  
302 ACTTGTGTGTATCTGTGAGACCCAGACTAGTGGCTCCAGAACTTAGAAAAAGCCAT 361  
383 ACTTGTGTGTATCTGTGAGACCCAGACTAGTGGCTCCAGAACTTAGAAAAAGCCAT 442  
362 CCAATTTATGTACCAAAACCTGACAGCAAGATGGGCTGAGAAAGTCACTCGAGAGCCAGT 421  
443 CCAATTTATGTACCAAAACCTGACAGCAAGATGGGCTGAGAAAGTCACTCGAGAGCCAGT 502  
422 GAGAATACCCCACTGAGAGAGGAGAGAAATCACTGTGATGCTGAGCCAGAAATTC 481  
503 GAGAATACCCCACTGAGAGAGGAGAGAAATCACTGTGATGCTGAGCCAGAAATTC 562  
482 TAAGATAGCCATCTGTGGTCTTGGCAAGCAGATTGGACTCTCCAGAAAGGCATTAC 541  
563 TAAGATAGCCATCTGTGGTCTTGGCAAGCAGATTGGACTCTCCAGAAAGGCATTAC 622  
542 AGAAGTCTGTGTGAGACCTTTTCATGATGACGAGAGAGGCTTAAAGAGCAAGAG 601  
623 AGAAGTCTGTGTGAGACCTTTTCATGATGACGAGAGAGGCTTAAAGAGCAAGAG 682  
602 GAAGATTGTTGTTATTAACAACCTTACATCACTACTCAAGAGCGGTGCAATCCGAG 661  
683 GAAGATTGTTGTTATTAACAACCTTACATCACTACTCAAGAGCGGTGCAATCCGAG 742  
662 GCAGAGGCGGCTGAGAGCTGCAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGG 721  
743 GCAGAGGCGGCTGAGAGCTGCAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGG 802  
722 CTCCTTCTCCATCTACAGATCTCTCACAGAGATATTCAGAAATACAGAGATGGCGCC 781  
803 CTCCTTCTCCATCTACAGATCTCTCACAGAGATATTCAGAAATACAGAGATGGCGCC 862  
782 AATTCACAGAGCTGTATTAAGGTGAGAGATGACAGAAATGATGTCAAGATGGCTTCA 841

Db 863 GATTCACAGAGCTGTATTAAGGTGAGAGATGACAGAAATGATGTCAAGATGGCTTCA 922  
Qy 842 TGGGATTAATATGTCTATTGAGCTTAAGAT-GGGGCGCAAGACCTTACCAGATCTGAT 900  
Db 923 TGGGATTAATATGTCTATTGAGCTTAAGATGGGGCGCAAGACCTTACCAGATCTGAT 982  
Qy 901 CCTTCAACACTGTAGCAGAGATCACTGGAGCAATATTCAGAAACAGATTGTACTG 956  
Db 983 CCTTCAACACTGTAGCAGAGATCACTGGAGCAATATTCAGAAACAGATTGTACTG 1038

RESULT 8  
LOCUS BX439466/c 1020 bp mRNA linear EST 15-MAY-2003  
DEFINITION BX439466 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE010Y006  
3-PRIME, mRNA sequence.  
ACCESSION BX439466  
VERSION BX439466.1 GI:30775754  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1020)  
L1, W.B., Gruber, C., Jesssee, J. and Polayres, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4663.f for  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DE010DE03NP1&cluster=4663.f. Contact :  
Peng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DE010DE03NP1.

# FEATURES

Location/Qualifiers  
1. 1020  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE010Y006"  
/issue\_type="PLACENTA"  
/clone\_id="Homo sapiens PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo (dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and EcoRV into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

BASE COUNT 272 a 224 c 223 g 266 t 35 others  
ORIGIN

Query Match 50.8%; Score 925.6; DB 13; Length 1020;  
Best Local Similarity 95.9%; Pred. No. 9.9e-166;  
Matches 925; Conservative 20; Mismatches 20; Indels 0; Gaps 0;

850 AATTTGATTCACCTTAAGATGGGGCAAGACCTTCCAGATGATGATCTCTTACA 909  
966 AATTTGATTCACCTTAAGATGGGGCAAGACCTTCCAGATGATGATCTCTTACA 907  
910 CTGTAGCAGATCACTGGAGCAAAATATTCAGAAACAGTTGTACTGTGAGTGAATC 969  
906 CTGTAGCAGATCACTGGAGCAAAATATTCAGAAACAGTTGTACTGTGAGTGAATC 847  
970 TGGACAGCTGGATGTTGGGAGGAGGCTCCATGATGATGAGCGGTGAGCCCTTATATCAT 1029  
846 TGGACAGCTGGATGTTGGGAGGAGGCTCCATGATGATGAGCGGTGAGCCCTTATATCAT 787  
1030 GGGAGACACTCTCACTTAATTAAGATCTTGGGGCTGGCTCCAAAGAGACTCTGGCGCTGG 1089

QY	Db	786	GGGAGCACTCTCACTTATTAAAGATCTTGGGCTGGCTGCACCAAGAGACTCTGGGGCTGG	727
QY	1090	TGCTCTGACATGACAGAAAGAACAGGTGAGTTGGTGGCTTCCAGTATTATCATGTTACACA	1144	
Db	726	TGCTCTGACATGACAGAAAGAACAGGTGAGTTGGTGGCTTCCAGTATTATCATGTTACACA	667	
QY	1150	AGGTAAATATTTTCCAACTACATCTGTGTGATGTGAGATCTTGACGGCAGAAACCTTCTTAA	1205	
Db	666	AGGTAAATATTTTCCAACTACATCTGTGTGATGTGAGATCTTGACGGCAGAAACCTTCTTAA	607	
QY	1210	CTGGGCTGCATTTCACTGGCAGATGAAAAAGGCGAGGGCCATCATYGGAGAGGTTATGAGCC	1268	
Db	606	CTGGGCTGCATTTCACTGGCAGATGAAAAAGGCGAGGGCCATCATYGGAGAGGTTATGAGCC	547	
QY	1270	TGCTGACAGCCCTCAATATCACTCAGAGTCTTGAGCCATGTGAGAGAGGACAGACATCAACT	1328	
Db	546	TGCTGACAGCCCTCAATATGACTCAGAGTCTTGAGCCATGTGAGAGAGGACAGACATCAACT	487	
QY	1330	TTTGGATCCAGCTGAGTGGCTGGAGCCAGTCTTCTTGATGACTTATTAACAATTTTCT	1389	
Db	486	TTTGGATCCAGCTGAGTGGCTGGAGCCAGTCTTCTTGATGACTTATTAACAATTTTCT	427	
QY	1390	TCTTCATCATCTCCACGGAGACACCATGACCTGTATGAGATGATCCAAAGCAGATGATGTTG	1448	
Db	426	TCTTCATCATCTCCACGGAGACACCATGACCTGTATGAGATGATCCAAAGCAGATGATGTTG	367	
QY	1450	CTGCTGCTGTGTTGGGCTGTGTTTCTTATGTTGTTGACAGACATGGAAGAAATGCTGCTTA	1508	
Db	366	CTGCTGCTGTGTTGGGCTGTGTTTCTTATGTTGTTGACAGACATGGAAGAAATGCTGCTTA	307	
QY	1510	GCTCTCAGAAACATGAGAAAGAAACGTTTATGCTTGGCCAGAACTCTGGGCTGTG	1566	
Db	306	GCTCTCAGAAACATGAGAAAGAAACGTTTATGCTTGGCCAGAACTCTGGGCTGTG	247	
QY	1570	CAACTTGGAAAACTCTTCAACAATTTGATGCAATTTATCTTCAAGACACAC	1628	
Db	246	CAACTTGGAAAACTCTTCAACAATTTGATGCAATTTATCTTCAAGACACAC	187	
QY	1630	TCTATTTCAATGCTTCTGTATTTATCTTCTTGTATCTTCCAAATCTCTGATCTTGA	1688	
Db	186	TCTATTTCAATGCTTCTGTATTTATCTTCTTGTATCTTCCAAATCTCTGATCTTGN	127	
QY	1690	AAAAGAAATCATTTCTCCCTCCCTCCACACATGAAATCAACATATGTGAGGATTTACA	1748	
Db	126	AAAAGAAATCATTTCTCCCTCCCTCCACACATGAAATCAACATATGTGAGGATTTACA	67	
QY	1750	GTGGGGGCAATTTCTTATATCACTCTTAAACATTTGTTCCACTTTTAAAGTAAACAC	1808	
Db	66	GTGGGGGCAATTTCTTATATCTCTTAAACATTTGTTCCACTTTTAAAGTAAACAM	7	
QY	1810	TTAAT 1814		
Db	6	TTTWT 2		
RESULT 9				
LOCUS	BX360506/c	1073 bp	mRNA	linear
DEFINITION	BX360506 Homo sapiens PLACENTA COR 25-NORMALIZED Homo sapiens cDNA			
ACCESSION	BX360506			
VERSION	BX360506.1			
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.			
TITLE	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
JOURNAL	Full-length cDNA libraries and normalization			
COMMENT	Unpublished			
	Contact: Genoscope			

FEATURES	Source
<p>Genoscope - Centre National de Sequencage            BP 191 91006 Evry cedex - France            Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr            Library was constructed by Life Technologies, a division of            Invitrogen. This sequence belongs to sequence cluster 4663.f For            more information about this cluster, see  <a href="http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1071AA10NP1&amp;cluster=4663.f">http://www.genoscope.cns.fr/</a>  <a href="http://fulllength.invitrogen.com">http://fulllength.invitrogen.com</a> URL :            Faraday Avenue Genoscope sequence ID : CS0D1071AA10NP1.</p>	
1. 1073	Location/Qualifiers
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="CS0D1071A13"	
/issue_type="PLACENTA COT 25-NORMALIZED"	
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"	
/note="1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."	
BASE COUNT	287 a 229 c 244 g 285 t 28 others
ORIGIN	
Query Match	48.3%; Score 894; DB 13; Length 1073;
Best Local Similarity	95.9%; Pred. No. 9.9e-160;
Matches	949; Conservative 10; Mismatches 27; Indels 4; Gaps 4;
QY	808 AAGATCGAAGAAATGATGTCAAGAAATGGCTTCTCATGTGGATCAAAATTTGTCTATTACGCTAA 867
DB	987 ACGGTGAARATSAATAATATGTCAAGATGCTTCATGATCAAAATGTCAATTAGCTAA 928
QY	868 AGATGGGGGCAAAACCTTACCAGATATCTGATTCCTTCAACACTGTAGACAGATATCACTG 927
DB	927 AGAT-GGGGCAAAACCTTACCAGATATCTGATTCCTTCAACACTGTAGACAGATATCACTG 870
QY	928 GGAGCAATATTCAGAAACAGGTTTACTGTGTCACTGTGAGACATCTGGACAGCTGGAGTTTG 987
DB	869 GGAGCAATATTCAGAAACAGGTTTACTGTGTCACTGTGAGACATCTGGACAGCTGGAGTTTG 810
QY	988 GGCAGGGTGCATATGATGATGATGGCGGTGAGAGCTTTATATCATGGGAAGCACTTCACTTA 1047
DB	809 GGCAGGGTGCATATGATGATGATGGCGGTGAGAGCTTTATATCATGGGAAGCACTTCACTTA 750
QY	1048 TTAAGAATCTTTGGGCTGGTGCAGAAAGAGACTGCGCGGTGGTCTGACATGCTGACAGAG 1107
DB	749 TTAAGAATCTT-GGCTGGTGCAGAAAGAGACTGCGCGGTGGTCTGACATGCTGACAGAG 691
QY	1108 AACAGGTGAGTTGGTGCTTCCAGATATTAATCATGTTAACACAGATAATATTTCAACT 1167
DB	690 AACAGGTGAGTTGGTGCTTCCAGATATTAATCATGTTAACACAGATAATATTTCAACT 631
QY	1168 ACACTCTGTGTATGAGATCTGACGCGAGAACTTCTTAACCATCTGGCTGCAATTCACTG 1227
DB	630 ACACTCTGTGTATGAGATCTGACGCGAGAACTTCTTAACCATCTGGCTGCAATTCACTG 571
QY	1228 GCACTGAAAAGG-CCAGGGCCATATATGAGAGAGTTATGAGCCGTGCGAGCCCCCTCAAT 1286
DB	570 GCACTGAAAAGGCCAGGGCCATATATGAGAGAGTTATGAGCCGTGCGAGCCCCCTCAAT 511
QY	1287 ATCACTCAGGCTCTGACGCGCATGAGAGAGGACACATCAACTTTTGGATCCAGCTTGA 1346
DB	510 ATCACTCAGGCTCTGACGCGCATGAGAGAGGACACATCAACTTTTGGATCCAGCTTGA 451
QY	1347 GTGCTGAGAGCCAGCTTAATTGATGACTTATACAGATATTTCTTCCATCACTCCAC 1406
DB	450 GTGCTGAGAGCCAGCTTAATTGATGACTTATACAGATATTTCTTCCATCACTCCAC 391
QY	1407 GGAAGACATGATCTGTATGATATCCAAAGCATATGATTTGCTGCTGCTTTGGGCT 1466
DB	390 GGAAGACATGATCTGTATGATATCCAAAGCATATGATTTGCTGCTGCTTTGGGCT 331

Qy	1467	GTGTGTTCTTAATGTTGTGCAGACATGGAAGAAATGCGCTAGTCTCTTAATAACAGTAA	1566
Db	330	GTGTGTTCTTAATGTTGTGCAGACATGGAAGAAATGCGCTAGTCTCTTAATAACAGTAA	271
Qy	1527	GAAGAAACGTTTTCATGCTTCTGGCCAGAAATCCTGGGTCGCAACTTTGGAAATCTCC	1586
Db	270	GAAGAAACGTTTTCATGCTTCTGGCCAGAAATCCTGGGTCGCAACTTTGGAAATCTCC	211
Qy	1587	TCTTCACTAACAATTCATCOAATTCATCTTCAAGCAACACTCTAATTCATGCTTTCT	1646
Db	210	TCTTCACTAACAATTCATCOAATTCATCTTCAAGCAACACTCTAATTCATGCTTTCT	151
Qy	1647	GTATATATCTTTCTTGATACCTTCCAAATTCCTGATCTTGAAGAAAGGAATCATTTCTCC	1706
Db	150	GTATATATCTTTCTTGATACCTTCCAAATTCCTGATCTTGAAGAAAGGAATCATTTCTCC	91
Qy	1707	CCTCCCTCCACACATAGAAATCAACATATGATGGAGATPACAGTGGGGGCAATTTCTTTA	1766
Db	90	CCTCCCTCCACACATAGAAATCAACATATGATGGAGATPACAGTGGGGGCAATTTCTTTA	31
Qy	1767	TATCACTCTTAAAAACATGTTTCACTT	1796
Db	30	TATCACTCTTAAAAACATGTTTCACTT	1

RESULT 10	1126 bp	mrna	linear	EST 05-MAY-2003
BX360507				
LOCUS				
DEFINITION	BX360507 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA			
	clone CS001071A19 5-PRIME, mRNA sequence.			

SOURCE ORGANISM	Homo sapiens (human)	Homo sapiens

REFERENCE 1 (bases 1 to 1126)

**JOURNAL** Unpublished  
**COMMENT** Contact: Genoscope

Genoscope - Centre National de Séquençage  
BP 191 91006 EVRY cedex - France  
Email: [secreter@genoscope.cns.fr](mailto:secreter@genoscope.cns.fr)  
Web: [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4663.f For  
more information about this cluster, see  
[http://www.genoscope.cns.fr/  
cg3-bin/cluster.cgi?seq=CS0D1071AA100P1&cluster=4663.f](http://www.genoscope.cns.fr/cg3-bin/cluster.cgi?seq=CS0D1071AA100P1&cluster=4663.f). Contact :  
Peng Liang Email : [liangpeng@life-techn.com](mailto:liangpeng@life-techn.com) URL :  
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>  
Paradey Avenue Genoscope sequence ID : CS0D1071AA100P1.

FEATURES	Location/Qualifiers
source	1. .1126

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1071YA19"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NctI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Nct I and cloned into the Nct I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized
BASE COUNT      311 a      249 c      298 g      246 t
ORIGIN            22 others

```

Query Match	47.1%	Score 872.4;	DB 13;	Length 1126;
Best Local Similarity	97.0%;	Pred. No. 1.2e-155;		
Matches 931; Conservative	9;	Mismatches 15;	Indels 5;	Gaps

QY	1	GGCAAGGCGCGAGCTTAAGTCCTGGAGCGCGCTCGCGCGCGCTACAGGCGCGCTTA	60
Db	112	GGCTAGGCGCGAGCTTAAGTCCTGGAGCGCGCTCGCGCGCGCTACAGGCGCGCTTA	171
QY	61	TCAGATTATCTTAAACAAGAAACCAACTGGAAAAAATGAAATTCCTTATCTTCGCAT	120
Db	172	TCAGATTATCTTAAACAAGAAACCAACTGGAAAAAATGAAATTCCTTATCTTCGCAT	231
QY	121	TTTTGGGTGGTGTACACCTTTTATCCCTGTGCTCTGGGAAAGCTATATTCAGAAATGGCA	180
Db	232	TTTTGGGTGGTGTACACCTTTTATCCCTGTGCTCTGGGAAAGCTATATTCAGAAATGGCA	291
QY	181	TCTTAAGAGGACCTTTGGAAGAAATAAAGAGAAATAGCAGCTGGAGATGTGTCTA	240
Db	292	TCTTAAGAGGACCTTTGGAAGAAATAAAGAGAAATAGCAGCTGGAGATGTGTCTA	351
QY	241	AAGCAATCATCAACTAGCTGTTATGTGTAAGCCCAAGACAGATCCTATAGCGATTTGG	300
Db	352	AAGCAATCATCAACTAGCTGTTATGTGTAAGCCCAAGACAGATCCTATAGCGATTTGG	411
QY	301	CACCTTCGTGTGATACGTGTGGACCCGACTGATGTGCTCCAGAAACCTTGAAGAAAGCCA	360
Db	412	CACCTTCGTGTGATACGTGTGGACCCGACTGATGTGCTCCAGAAACCTTGAAGAAAGCCA	471
QY	361	TCCAAATTTATGTACAAACCTGACAGACATGGCGCTGAGAAAGTTCACCTGGAGCGAG	420
Db	472	TCCAAATTTATGTACAAACCTGACAGACATGGCGCTGAGAAAGTTCACCTGGAGCGAG	531
QY	421	TGAGAAATACCCCACTGGGAGAGGGGAGAGAAATCAGCTGTGATGCTGGAGCCAAAGATTC	480
Db	532	TGAGAAATACCCCACTGGGAGAGGGGAGAGAAATCAGCTGTGATGCTGGAGCCAAAGATTC	591
QY	481	ATAAGATATGCCATCTCTGGGTCTTGGGACGACGATTTGGGAATCTCTCCAGAAAGCATTACAG	540
Db	592	ATAAGATATGCCATCTCTGGGTCTTGGGACGACGATTTGGGAATCTCTCCAGAAAGCATTACAG	651
QY	541	CAGAAGTTCGTGGTGGTGAACCTCTTCGATGAACTGACAGAGAAAGGGGCTCAGAAAGCAAGAG	600
Db	652	CAGAAGTTCGTGGTGGTGAACCTCTTCGATGAACTGACAGAGAAAGGGGCTCAGAAAGCAAGAG	711
QY	601	GGAAGATTGTTGTTTATTAACCAACCTTACATCAACTACAGAGACGGTGCATATCCGAA	660
Db	712	GGAAGATTGTTGTTTATTAACCAACCTTACATCAACTACAGAGACGGTGCATATCCGAA	771
QY	661	CGCAGGGGGCGGTGGAAAGCTGCCAAGGTGGGGCTTTGGCATCTCTCATTTGCATCCGTGG	720
Db	772	CGCAGGGGGCGGTGGAAAGCTGCCAAGGTGGGGCTTTGGCATCTCTCATTTGCATCCGTGG	831
QY	721	CCTCCTTCTCGATTAACGTCCTCACACAGGTATTGAGAAATACAGAGATGGCGTGGCCA	780
Db	832	CCTCCTTCTCGATTAACGTCCTCACACAGGTATTGAGAAATACAGAGATGGCGTGGCCA	891
QY	781	AAATTCACAACGCTGTATTACGGT-GGAAGATGCAGAAATGATGTCAAGAAATGGCTTCT	839
Db	892	AAATTCACAACGCTGTATTACGGT-GGAAGATGCAGAAATGATGTCAAGAAATGGCTTCT	951
QY	840	CA-TGGGATCAAAATTTGTCAATTCAGCTTAAAGATGGGGGCAAAAGCTTAAACCAATATCTGA	898
Db	952	CATTGGGATCAAAATTTGTCAATTCAGCTTAAAGATGGGGG- AARMCATACCAGATCTCTGA	1010
QY	899	TTCCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATTCAGAAACGGTTGTACTGCT	958
Db	1011	TTCCCTT-ACACGTATAGCAR-ATCACTGTGGAGCAAAATCCARMAAGGTACTACGTAGCT	1068

RESULT	11
BQ878966	
LOCUS	
DEFINITION	BQ878966 909 bp mRNA linear EST 16-AUG-2002
ACCESSION	AGNCOCURT_8183107 lupsk1_dorsal_root ganglion Homo sapiens cDNA
VERSION	clone IMAGE:6184129 5', mRNA sequence.
KEYWORDS	BQ878966 BQ878966.1 GI:22270974
	EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 909)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-remail.nih.gov  
 Tissue Procurement: Dr. James R. Lupski  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.llnl.gov  
 Plate: L1M13572 row: n column: 02  
 High quality sequence stop: 696.  
 Location/Qualifiers  
 1. 909  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6184129"  
 /sex="male"  
 /tissue\_type="dorsal root ganglia"  
 /dev\_stage="adult, 36 yr"  
 /lab\_host="DH10B"  
 /clone\_lib="Lupski dorsal root ganglion"  
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:  
 NotI; Site 2: SalI; cDNA made by oligo-dT priming.  
 Directionally cloned using the following adaptors:  
 5'-GTGAGTCTAGATCGAGCGCGCCCT(15)-3', Size selected >  
 1 kb for average insert length 1.7 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine) and is available through Life  
 Technologies."

BASE COUNT 273 a 202 c 226 g 207 t 1 others  
 ORIGIN

Query Match 45 8%; Score 847.8; DB 13; Length 909;  
 Best Local Similarity 98.8%; Pred. No. 6.2e-151;  
 Matches 885; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

Qy 37 GTCCGCCGCTCAGAGCCGCTATCAGATTATCTTAACAAGAAACCACT-GGAAAAA 95  
 Db 1 GTCCGCCGCTCAGAGCCGCTATCAGATTATCTTAACAAGAAACCACTGGGAAAAA 60  
 Qy 96 AAATGAATTCCTTATCTTGCGATTTTGGGTGTTACCTTTATCCCTGTCTCT 155  
 Db 61 AAATGAATTCCTTATCTTGCGATTTTGGGTGTTACCTTTATCCCTGTCTCT 120  
 Qy 156 GGAAGAGTATATGAAAGATGCGATCTTAAGAGACTTTTGAAGAAATTAAGAAATA 215  
 Db 121 GGAAGAGTATATGAAAGATGCGATCTTAAGAGACTTTTGAAGAAATTAAGAAATA 180  
 Qy 216 ATAGCAGCTGTGAGATGTTGCTTAAGCAATCATCACTAGCTGTTTATGTTAAACC 275  
 Db 181 ATAGCAGCTGTGAGATGTTGCTTAAGCAATCATCACTAGCTGTTTATGTTAAACC 240  
 Qy 276 CAGAAAGATCTTATAGCGATGCGACTTCTGTTGTAAGTGTGAACCAAGACTGAGT 335  
 Db 241 CAGAAAGATCTTATAGCGATGCGACTTCTGTTGTAAGTGTGAACCAAGACTGAGT 300  
 Qy 336 GGGTTCAGAACTTGAAGAAAGCCATCAAAATTAATGTTCAAAAACCTGAGAGATGGG 395  
 Db 301 GGGTTCAGAACTTGAAGAAAGCCATCAAAATTAATGTTCAAAAACCTGAGAGATGGG 360  
 Qy 396 CTGGAAGAGTTTCACTGAGAGCCAGTGAAGATACCCCTGGAGAGGGGAGAGATCA 455

Db 361 CTAGAGAAAGTTTCACTGAGAGCCAGTGAATATACCCCACTGGAGAGGAGAGAAATCA 420  
 Qy 456 GCTGTATGCTGAGAGCCAGAAATTCATTAAGATAGCCATCTTGCTTTGGACAGAGATT 515  
 Db 421 GCTGTATGCTGAGAGCCAGAAATTCATTAAGATAGCCATCTTGCTTTGGACAGAGATT 480  
 Qy 516 GGGACTCTCTCAGAAAGGCTTTACAGCAGAGTTCTGTGTGACCTCTTTGATGAATCTG 575  
 Db 481 GGGACTCTCTCAGAAAGGCTTTACAGCAGAGTTCTGTGTGACCTCTTTGATGAATCTG 540  
 Qy 576 CAGAGAGAGGCTCAGAGAGAGGAGAGAGATGTTTATTAACAACCTTACATCAAC 635  
 Db 541 CAGAGAGAGGCTCAGAGAGAGGAGAGAGATGTTTATTAACAACCTTACATCAAC 600  
 Qy 636 TACTCAAGACGGTGCATTCAGAACGAGAGGCGGTGGAAGCTGCAGAGTGGGGCT 695  
 Db 601 TACTCAAGACGGTGCATTCAGAACGAGAGGCGGTGGAAGCTGCAGAGTGGGGCT 660  
 Qy 696 TTGGCATCTCTCATTCGATCCGTGGCTCCTCTCTCATCTTACAGTCTTACACAGTATT 755  
 Db 661 TTGGCATCTCTCATTCGATCCGTGGCTCCTCTCTCATCTTACAGTCTTACACAGTATT 720  
 Qy 756 CAGGAATACAGAGATGCGGCGCCCAAAATTCACAGCTGTATTAAGGTGAAGATCA 815  
 Db 721 CAGGAATACAGAGATGCGGCGCCCAAAATTCACAGCTGTATTAAGGTGAAGATCA 780  
 Qy 816 GAATGATGTCAGAAATGCGCTTCTCATAGGATCAAAATTTGCTTCATTCGTTAAAGAT 874  
 Db 781 GAATGATGTCAGAAATGCGCTTCTCATAGGATCAAAATTTGCTTCATTCGTTAAAGAT 840  
 Qy 875 GGCAGAGCTTACCCAGATAGTCTTCTTCAACT-GTAGCAGAGATCACTGGG 929  
 Db 841 GGCAGAGCTTACCCAGATAGTCTTCTTCAACTGATCTTTCACACTGATGAGATCACTGGG 896

RESULT 12  
 BX416895/c  
 LOCUS  
 DEFINITION  
 BX416895 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DB001YF04  
 3-PRIME, mRNA sequence.  
 ACCESSION  
 BX416895  
 VERSION  
 BX416895.1 GI:30650311  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1003)  
 AUTHORS Li, W.-B., Gruber, C., Jesse, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seq@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 4663.f For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DB001D02NP1cluster=4663.f. Contact :  
 Feng Liang Email : fliang@life.techn.com URL :  
 http://fulllength.invitrogen.com/InvitrogenCorporation1600  
 Faraday Avenue genoscope sequence ID : CS0DB001D02NP1.

FEATURES  
 Location/Qualifiers  
 1. 1003  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DB001YF04"  
 /tissue\_type="PLACENTA"  
 /clone\_lib="Homo sapiens PLACENTA"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo (dT) primer. Five prime end enriched,



```

QY 301 CACTTCTGTTGATCTGTTGGACCCAGACTGAGTGGCTCAAGAACTAGAAAAAGCCA 360
DB 457 CACTTCTGTTGATCTGTTGGACCCAGACTGAGTGGCTCAAGAACTAGAAAAAGCCA 516
QY 361 TCCTAAATTTATGACCAAACTGCGACGAAAGATGGCTGAGAAAGTTCACTGAGCCAG 420
DB 517 TCCTAAATTTATGACCAAACTGCGACGAAAGATGGCTGAGAAAGTTCACTGAGCCAG 576
QY 421 TGAGATATCCCACTGGGAGAGGGGAGAGATACAGCTGTGATGCTGGAGCCCAAAATTC 480
DB 577 TGAGATATCCCACTGGGAGAGGGGAGAGATACAGCTGTGATGCTGGAGCCCAAAATTC 636
QY 481 ATAGATAGCCATCTGAGTCTTGGACAGACATGGGACTCTCTCAGAAAGCATTAACAG 540
DB 637 ATAGATAGCCATCTGAGTCTTGGACAGACATGGGACTCTCTCAGAAAGCATTAACAG 696
QY 541 CAGAAATTTGTTGATGACCTCTTTCGATGAACTGCAAGAAAGGCTTCAGAAAGAGAG 600
DB 697 CAGAAATTTGTTGATGACCTCTTTCGATGAACTGCAAGAAAGGCTTCAGAAAGAGAG 756
QY 601 GGAAGATTTGTTTATATACCAACTTATGATCACTACAGAGAGCGGTGCAATACCGAA 660
DB 757 GGAAGATTTGTTTATATACCAACTTATGATCACTACAGAGAGCGGTGCAATACCGAA 816
QY 661 CGCAGGGGGCGGTGAGAGCTGCGCAAGG-TGGGGGCTTTGGCATCTCTCATTCGATCCG 719
DB 817 CGCAGGGGGCGGTGAGAGCTGCGCAAGG-TGGGGGCTTTGGCATCTCTCATTCGATCCG 876
QY 720 GCTCTCTTTCATCTACAGTCTCTCAACAGATTTACAGAAATACAGAGTGGCTGCC 779
DB 877 GCTCTCTTTCATCTACAGTCTCTCAACAGATTTACAGAAATACAGAGTGGCTGCC 936
QY 780 AAAATTTCAACAGCCGTATTAAGGTGAGAGATGCAAGAAATGATGCAAGATGCTCT 839
DB 937 AAGATTTCAACAGCCGTATTAAGGTGAGAGATGCAAGAAATGATGCAAGATGCTCT 996
QY 840 CATGGATCAAAATTTG 855
DB 997 CATGGATCAAAATTTG 1012

```

```

RESULT 14
LOCUS BX335995 919 bp mRNA linear EST 02-MAY-2003
DEFINITION BX335995 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
ACCESSION BX335995
VERSION BX335995.1 GI:30339459
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 919)
li.w.b., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgl-bin/cluster.cgi?seq=CS0D1022DE08QPlcluster=4663.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1022DE08QPl.
Location/Qualifiers
1..919
/organism="Homo sapiens"

```

```

/mot type="mRNA"
/cdb xref="taxon:9606"
/clone="CS0D1022YJ16"
/issue type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/clone="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 273 a 206 c 226 g 208 t 6 others
ORIGIN

```

```

Query Match 45.3%; Score 838.6; DB 13; Length 919;
Best Local Similarity 96.9%; Pred. No. 3,4e-149;
Matches 861; Conservative 23; Mismatches 1; Gaps 1;

```

```

QY 1 GCTAGGCCGCGACCTTAGTCTCTGAGAGCCGCTCCGCTGCGCGCTGAGAGCCGCTTA 60
DB 28 GGTGTATACCGGCTCCGGAATTCGGGAGTCCGCTCCGCTGCGCGCTGAGAGCCGCTTA 87
QY 61 TCAGATTTATCTTAACAGAAACCAACTGAAAGAAATTAATTCCTTATCTTCCAT 120
DB 88 TCAGATTTATCTTAACAGAAACCAACTGAAAGAAATTAATTCCTTATCTTCCAT 147
QY 121 TTTTCGGTGTGTACCTTTTATCCCTGCTCTGAGAAAGCTATATGCAAGATGCA 180
DB 148 TTTTCGGTGTGTACCTTTTATCCCTGCTCTGAGAAAGCTATATGCAAGATGCA 207
QY 181 TCTCTAAGAGACCTTTTGAAGAAATTAAGAAATGACCAAGCTGTGAGATGTTGCTA 240
DB 208 TCTCTAAGAGACCTTTTGAAGAAATTAAGAAATGACCAAGCTGTGAGATGTTGCTA 267
QY 241 AAGCAATCATCAACTGCTGTTTATGTTAAAGCCGAACAGATCTTATAGAGATTTG 300
DB 268 AAGCAATCATCAACTGCTGTTTATGTTAAAGCCGAACAGATCTTATAGAGATTTG 327
QY 301 CACTTCTGTTGATCTGTTGACCCAGACTGAGTGGCTCCAGAACTAGAAAAAGCCA 360
DB 328 CACTTCTGTTGATCTGTTGACCCAGACTGAGTGGCTCCAGAACTAGAAAAAGCCA 387
QY 361 TCCTAAATTTATGACCAAACTGCGACGAAAGATGGCTGAGAAAGTTCACTGAGCCAG 420
DB 388 TCCTAAATTTATGACCAAACTGCGACGAAAGATGGCTGAGAAAGTTCACTGAGCCAG 447
QY 421 TGAGATATCCCACTGGGAGAGGGGAGAGATACAGCTGTGATGCTGGAGCCCAAAATTC 480
DB 448 TGAGATATCCCACTGGGAGAGGGGAGAGATACAGCTGTGATGCTGGAGCCCAAAATTC 507
QY 481 ATAGATAGCCATCTGAGTCTTGGACAGACATGGGACTCTCTCAGAAAGCATTAACAG 540
DB 508 ATAGATAGCCATCTGAGTCTTGGACAGACATGGGACTCTCTCAGAAAGCATTAACAG 567
QY 541 CAGAAATTTGTTGATGACCTCTTTCGATGAACTGCAAGAAAGGCTTCAGAAAGAGAG 600
DB 568 CAGAAATTTGTTGATGACCTCTTTCGATGAACTGCAAGAAAGGCTTCAGAAAGAGAG 627
QY 601 GGAAGATTTGTTTATATACCAACTTATGATCACTACAGAGAGCGGTGCAATACCGAA 660
DB 628 GGAAGATTTGTTTATATACCAACTTATGATCACTACAGAGAGCGGTGCAATACCGAA 687
QY 661 CGCAGGGGGCGGTGAGAGCTGCGCAAGG-TGGGGGCTTTGGCATCTCTCATTCGATCCG 720
DB 688 CGCAGGGGGCGGTGAGAGCTGCGCAAGG-TGGGGGCTTTGGCATCTCTCATTCGATCCG 747
QY 721 CCTCTCTTTCATCTACAGTCTCTCAACAGATTTACAGAAATACAGAGATGGCTGCCA 780
DB 748 CCTCTCTTTCATCTACAGTCTCTCAACAGATTTACAGAAATACAGAGATGGCTGCCA 807
QY 781 AAATTTCAACAGCTGTATTAAGGTGAGAGATGCAAGAAATGATGCAAGATGCTCTC 840
DB 808 AGATTTCAACAGCTGTATTAAGGTGAGAGATGCAAGAAATGATGCAAGATGCTCTC 867
QY 841 AT-GGATCAAAATTTGATTCATCAAGTAAAGATGGGGCAAGACTTACC 888

```





**THIS PAGE BLANK (USPTO)**



```

OY 301 CACTTCGTGATGATCTGTGGAGCCAGAGCTAGTGGCTCCAGAAAGCTAGAAAAAGCA 360
Db 301 CACTTCGTGATGATCTGTGGAGCCAGAGCTAGTGGCTCCAGAAAGCTAGAAAAAGCA 360
OY 361 TCCTCAATTTATGATACCAAAAGCTGAGCAAGATGGGCTGAGAAAGTTCACTGAGGCA 420
Db 361 TCCTCAATTTATGATACCAAAAGCTGAGCAAGATGGGCTGAGAAAGTTCACTGAGGCA 420
OY 421 TGAAGATTCCTGAGGAGGAGGAGGAGAAATGAGTGTGATGTGTGAGGCAAGAAATTC 480
Db 421 TGAAGATTCCTGAGGAGGAGGAGGAGAAATGAGTGTGATGTGTGAGGCAAGAAATTC 480
OY 481 ATAGATGATGATCTGAGGATCTTGGCAGAGATGGAGCTCTCCAGAAAGCATTTAG 540
Db 481 ATAGATGATGATCTGAGGATCTTGGCAGAGATGGAGCTCTCCAGAAAGCATTTAG 540
OY 541 CAGAAATTCGTGAGGATCTTGGCAGAGATGGAGCTCTCCAGAAAGCATTTAG 600
Db 541 CAGAAATTCGTGAGGATCTTGGCAGAGATGGAGCTCTCCAGAAAGCATTTAG 600
OY 601 GGAAGATTCGTGATTTATTAACCAAGCTTATCACTCAAGAGCGGTGCAATTCGAA 660
Db 601 GGAAGATTCGTGATTTATTAACCAAGCTTATCACTCAAGAGCGGTGCAATTCGAA 660
OY 661 CGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
Db 661 CGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
OY 721 CTTCTCTCTCACTACAGTCTCTCAAGAGGATTTCAAGAAATTCAGAGATGGCGTCCA 780
Db 721 CTTCTCTCTCACTACAGTCTCTCAAGAGGATTTCAAGAAATTCAGAGATGGCGTCCA 780
OY 781 AAATTCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Db 781 AAATTCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
OY 841 ATGGATCAAAATTTGTCACTTCACTTCAAGATGGGAGGAGGAGGAGGAGGAGGAGGAG 900
Db 841 ATGGATCAAAATTTGTCACTTCACTTCAAGATGGGAGGAGGAGGAGGAGGAGGAGGAG 900
OY 901 CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAAGGTTGATCTGATCA 960
Db 901 CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAAGGTTGATCTGATCA 960
OY 961 GTGACATCTGAGCAGCTGGAGATGTTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
Db 961 GTGACATCTGAGCAGCTGGAGATGTTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
OY 1021 TTAATTCATGAGGAGCACTCTCACTTATTAAGATCTTGGGCTGGCTCCAAAGAGGAGCTC 1080
Db 1021 TTAATTCATGAGGAGCACTCTCACTTATTAAGATCTTGGGCTGGCTCCAAAGAGGAGCTC 1080
OY 1081 TGGGCTGGTGTCTGAGCTGAGCAAGAAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db 1081 TGGGCTGGTGTCTGAGCTGAGCAAGAAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
OY 1141 AGTTACACAAGGATTAATTTTCCAACTAGCTGTGATGAGGAGGAGGAGGAGGAGGAGGAG 1200
Db 1141 AGTTACACAAGGATTAATTTTCCAACTAGCTGTGATGAGGAGGAGGAGGAGGAGGAGGAG 1200
OY 1201 TCTTACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Db 1201 TCTTACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
OY 1261 TTAATGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Db 1261 TTAATGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
OY 1321 ACATTAACCTTTTGAATCCAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
Db 1321 ACATTAACCTTTTGAATCCAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
OY 1381 AGTATTTCTTCTTCATCACTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440

```

```

Db 1381 AGTATTTCTTCTTCATCACTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
OY 1441 TGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 1441 TGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
OY 1501 TGTGCTGAGGATCTTGAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
Db 1501 TGTGCTGAGGATCTTGAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
OY 1561 CTGGGCTGCAACTTTGAGAAAGCTCTCTTCACTAATCAATTTCAATTTCAATTTCAATTTCA 1620
Db 1561 CTGGGCTGCAACTTTGAGAAAGCTCTCTTCACTAATCAATTTCAATTTCAATTTCAATTTCA 1620
OY 1621 AAGCACAACCTTATTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Db 1621 AAGCACAACCTTATTTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
OY 1681 GATTCCTGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
Db 1681 GATTCCTGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
OY 1741 GGAATTAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
Db 1741 GGAATTAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
OY 1801 AGTAAACACTTAAATTTTGGAGAGATCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1851
Db 1801 AGTAAACACTTAAATTTTGGAGAGATCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1851

```

## RESULT 2

US-09-482-273-95  
; Sequence 95, Application US/09482273

; Patent No. 6534631

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OR INVENTION: 71 Human Secreted Proteins

; FILE REFERENCE: P2030P1

; CURRENT APPLICATION NUMBER: US/09/482,273

; EARLIER FILING DATE: 2000-01-13

; EARLIER APPLICATION NUMBER: PCT/US99/15849

; EARLIER FILING DATE: 1999-07-14

; EARLIER APPLICATION NUMBER: 60/092,921

; EARLIER FILING DATE: 1998-07-15

; EARLIER APPLICATION NUMBER: 60/092,956

; EARLIER FILING DATE: 1998-07-15

; NUMBER OF SEQ ID NOS: 267

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 95

; LENGTH: 1134

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-482-273-95

Query Match 59.3%; Score 1098.4; DB 4; Length 1134;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1121; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```

OY 729 TCCTATTCACAGTCTCTCAACAGGATTTCAAGAAATTCAGAGATGGCGGCCAAATTC 788
Db 1 TCCTATTCACAGTCTCTCAACAGGATTTCAAGAAATTCAGAGATGGCGGCCAAATTC 788
OY 789 ACAGCTGTATTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 848
Db 61 ACAGCTGTATTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 848
OY 849 AAAATTTGTCACTTACGCTTAAAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 908
Db 121 AAAATTTGTCACTTACGCTTAAAGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 908

```

RESULT 3  
US-08-232-463-14/C  
; Sequence 14, Application US/0832463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DONNER, F.  
; APPLICANT: SCHEIFFLINGER, F.

```

1  APPLICANT: FALKNER, F. G.
2  TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
3  NUMBER OF SEQUENCES: 52
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE: Foley & Lardner
6  STREET: 1800 Diagonal Road, Suite 500
7  CITY: Alexandria
8  STATE: VA
9  COUNTRY: USA
10 ZIP: 22313-0299
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: Floppy disk
14 COMPUTER: IBM PC compatible
15 OPERATING SYSTEM: PC-DOS/MS-DOS
16 SOFTWARE: Patentn Release #1.0, Version #1.25
17
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/08/232,463
20 FILING DATE:
21
22 CLASSIFICATION: 435
23
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US/07/935,313
26 FILING DATE:
27
28 APPLICATION NUMBER: EP 91 114 300.6
29 FILING DATE: 26-AUG-1991
30 ATTORNEY/AGENT INFORMATION:
31
32 NAME: BENT, Stephen A.
33 REGISTRATION NUMBER: 29,768
34 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: (703) 836-9300
37 TELEFAX: (703) 683-4109
38
39 TELEX: 899149
40
41 INFORMATION FOR SEQ ID NO: 14:
42
43 SEQUENCE CHARACTERISTICS:
44 LENGTH: 7218 base pairs
45 TYPE: nucleic acid
46 STRANDEDNESS: single
47 TOPOLOGY: linear
48
49 IMMEDIATE SOURCE:
50 CLONE: pTZgpc-F18
51
52 US-08-232-463-14
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
95
```

Oy	567	GATGA	CTGC	AGAGA	GGGCT	CA	590
Db	1051	GACCTG	CAGC	CAAGC	TCCGA	TTA	1028

RESULT 4  
US-08-48

```

1Sequence 50, Application US/08487001A
2Patent No. 5795862
3
4GENERAL INFORMATION:
5APPLICANT: FRANK, GLENN R.
6APPLICANT: HUNTER, SHIRLEY WU
7APPLICANT: WALTENFELS, LYNDA
8TITLE OF INVENTION: NOVEL ECOPARASITE SALIVA
9TITLE OF INVENTION: PROTEINS AND APPARATUS TO COLLECT SUCH PROTEINS
10NUMBER OF SEQUENCES: 54
11CORRESPONDENCE ADDRESS:
12ADDRESSEE: Sheridan Ross & McIntosh
13STREET: 1700 Lincoln Street, Suite 3500
14CITY: Denver
15STATE: Colorado
16COUNTRY: U.S.A.
17ZIP: 80203
18
19COMPUTER READABLE FORM:
20MEDIUM TYPE: Floppy disk
21COMPUTER: IBM PC compatible
22OPERATING SYSTEM: PC-DOS/MS-DOS
23SOFTWARE: PatentIn Release #1.0, Version #1.25
24CURRENT APPLICATION DATA:
25APPLICATION NUMBER: US/08/487,001A
26FILING DATE: 07-JUN-1995
27CLASSIFICATION: 424
28ATTORNEY/AGENT INFORMATION:
29NAME: Verser, Carol Talkington
30REGISTRATION NUMBER: 37,459
31REFERENCE/DOCKET NUMBER: 2618-17-C2
32TELECOMMUNICATION INFORMATION:
33TELEPHONE: (303) 863-9700
34TELEFAX: (303) 863-0223
35INFORMATION FOR SEQ ID NO: 50:
36SEQUENCE CHARACTERISTICS:
37LENGTH: 646 base pairs
38TYPE: nucleic acid
39STRANDEDNESS: single
40TOPOLOGY: linear
41MOLECULE TYPE: CDNA
42FEATURE:
43NAME/KEY: CDS
44LOCATION: 3..519
45US-08-487-001A-50

```

```

2.2%: Score 40; DB 1; Length 646;
Best Local Similarity 63.5%; Pred. No. 0.042;
Matches 61; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1756 GCATTCTTTATACCTCTTAAAAACATGTTCCACTTAAAGTAACCTTAATA 1815E
D8 551 GCTATTTTCTGTAAAAACATATATAAGCATTTTAACTGTTGTACGATATATACATAATA 610
QY 1816 AATTTTGAAGATCTCTGAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTT 1851
D8 611 AATTGCTACATTGCTCTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTT 646

RESULT 5
US-08-630-822A-50
; Sequence 50, Application US/08630822A
; Patent No. 5840695
; GENERAL INFORMATION:
; APPLICANT: FRANK, GLENN R.
; APPLICANT: HUNTER, SHIRLEY WU
; APPLICANT: WALLENFELS, LYNDA

```

TITLE OF INVENTION: NOVEL ECOPARASITE SALIVA PROTEINS  
 TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS  
 NUMBER OF SEQUENCES: 107  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sheridan Ross P.C.  
 STREET: 1700 Lincoln Street, Suite 3500  
 CITY: Denver  
 STATE: Colorado  
 COUNTRY: U.S.A.  
 ZIP: 80203  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/630,822A  
 FILING DATE: 11-APR-1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: CONNELL, GARY J.  
 REGISTRATION NUMBER: 32,020  
 REFERENCE/DOCKET NUMBER: 2618-17-C3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (303) 863-9700  
 TELEFAX: (303) 863-0223  
 INFORMATION FOR SEQ ID NO: 50:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 646 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 3..519  
 US-08-630-822A-50

[illegible]

RESULT 6  
 US-09-005-069-50  
 Sequence 50 Application US/0905069  
 Patent No. 5832470  
 GENERAL INFORMATION:  
 APPLICANT: FRANK, GLENN R.  
 APPLICANT: HUNTER, SHIRLEY WU  
 APPLICANT: MALLENFELS, LYNDA  
 TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS  
 TITLE OF INVENTION: AND APPARATUS TO COLLECT SUCH PROTEINS  
 NUMBER OF SEQUENCES: 107  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sheridan Rose P. C.  
 STREET: 1700 Lincoln Street, Suite 3500  
 CITY: Denver  
 STATE: Colorado  
 COUNTRY: U.S.A.  
 ZIP: 80203  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/005,069  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/630,822  
FILING DATE: 11-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: CONNELL, GARY J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-17-C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 646 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..519  
US-09-005-069-50

Query Match 2.2%; Score 40; DB 2; Length 646;  
Best Local Similarity 63.5%; Pred. No. 0.042;  
Matches 61; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 1756 GCATTTCTTATACACCTCTAAACATGTTCCCTTAAAGTAACCTTATA 1815  
Db 551 GCTATTTCTTAAACCATTAAGCTATTTTAACTTGTACAGTATACATA 610  
Qy 1816 AATTTTGAAGATCTGAAAAA 1851  
Db 611 AATTGCTACATTTGCTTAAAAA 646

RESULT 7  
US-08-232-463-14  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEFFLINGER, F.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-9300  
TELEFAX: (703) 683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZgpc-Fls  
US-08-232-463-14

Query Match 2.2%; Score 40; DB 1; Length 7218;  
Best Local Similarity 4.7%; Pred. No. 0.22;  
Matches 19; Conservative 211; Mismatches 176; Indels 0; Gaps 0;

Qy 1354 GAGCACTCTACTGATGACTTATACAGTATTTCTTCCATCACTCCACGAGACA 1413  
Db 1064 GATTT 1123  
Qy 1414 CCATGACTGTCATGATCAAGACATGATGTCGCTGTTGGCTGTGTTT 1473  
Db 1124 TTT 1183  
Qy 1474 CTATGTTGTGACAGATGAGAAATGCTGCTAGCTCTAGAAACAGTAAAGAA 1533  
Db 1184 TTT 1243  
Qy 1534 ACGTTTCACTCTGCGCAGCAATCTGCTGCTGCACTTGGAATCTCTTAC 1593  
Db 1244 TTT 1303  
Qy 1594 ATAACTATTCATCATTCATCTCAAGACACTCATTCATCTTCTGTATTA 1653  
Db 1304 TTT 1363  
Qy 1654 TCTTCTGACTTCCAAATCTCTGATTTAGAAAAAGATCATCTCCCTCCT 1713  
Db 1364 TTT 1423  
Qy 1714 CCCACCATGATCATCATATGATGATGATGATGATGATGATGATGAT 1759  
Db 1424 TTTTTTTTTTTTACCAATCTCTATCTTTAACTACTTGAT 1469

RESULT 8  
US-08-268-797-1/C  
Sequence 1, Application US/08268797  
Patent No. 5599788  
GENERAL INFORMATION:  
APPLICANT: Purchio, Anthony F.  
APPLICANT: Lebaron, Richard  
TITLE OF INVENTION: Factor to Grow Tissue Ex Vivo  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Knobe, Martens, Olson and Bear  
STREET: 620 Newport Center Drive, Sixteenth Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/268,797  
FILING DATE:

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned A.  
REGISTRATION NUMBER: 29,655  
TELEPHONE/DOCKET NUMBER: TISSUE.001A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2049 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-268-797-1

Query Match 2.1%; Score 38; DB 1; Length 2049;  
Best Local Similarity 47.8%; Pred. No. 0.39;  
Matches 110; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

Qy 1295 GGTCTGAGCCATGAGAGGAGAGACATCACTTTGGATCCAGCTGAGTGCCTGG 1354  
Db 1900 GCTGAGAGAACTGTTGATGATGACATGAGACGACGCAATTTGTGGCATGATGTCAGGCTGG 1841  
Qy 1355 AGCCAGTCTACTTGATGATGATTTATCAAGTATTTCTTCCATCACTCCCGAGAGACAC 1414  
Db 1840 CAACAGGCTCTTGTGATGACACTCACCACTGTTTTCAGAGCTGATCCAGCTTGTAC 1781  
Qy 1415 CATGACTGTCAATGATCCAAAGAGATGATGCTGCTGCTGTTGGGCTGTTGTTTC 1474  
Db 1780 CTTGAGAGACTTTAGCCGACCAAGGCCCCGATGCTTCCGCTAACAGAGATTTATCAGC 1721  
Qy 1475 TTATGTTGTGAGACATGAGAAATGCTGCTAGTCTTGAAGAACT 1524  
Db 1720 CAATGTGATTTTCAAGATGTTGGCAAGTCTTGGCATCTCCCAAGAGT 1671

RESULT 9  
PCT-US95-08414-1/c  
Sequence 1, Application PC/TUS9508414  
GENERAL INFORMATION:  
APPLICANT: Purchio, Anthony F.  
TITLE OF INVENTION: Factor to Grow Tissue Ex Vivo  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson and Bear  
STREET: 620 Newport Center Drive, Sixteenth Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/08414  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/268,797  
FILING DATE: July 1, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: TISSUE.001A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2049 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
PCT-US95-08414-1

Query Match 2.1%; Score 38; DB 5; Length 2049;  
Best Local Similarity 47.8%; Pred. No. 0.39;  
Matches 110; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

Qy 1295 GGTCTGAGCCATGAGAGGAGAGACATCACTTTGGATCCAGCTGAGTGCCTGG 1354  
Db 1900 GCTGAGAGAACTGTTGATGATGACATGAGACGACGCAATTTGTGGCATGATGTCAGGCTGG 1841  
Qy 1355 AGCCAGTCTACTTGATGATGATTTATCAAGTATTTCTTCCATCACTCCCGAGAGACAC 1414  
Db 1840 CAACAGGCTCTTGTGATGACACTCACCACTGTTTTCAGAGCTGATCCAGCTTGTAC 1781  
Qy 1415 CATGACTGTCAATGATCCAAAGAGATGATGCTGCTGCTGTTGGGCTGTTGTTTC 1474  
Db 1780 CTTGAGAGACTTTAGCCGACCAAGGCCCCGATGCTTCCGCTAACAGAGATTTATCAGC 1721  
Qy 1475 TTATGTTGTGAGACATGAGAAATGCTGCTAGTCTTGAAGAACT 1524  
Db 1720 CAATGTGATTTTCAAGATGTTGGCAAGTCTTGGCATCTCCCAAGAGT 1671

RESULT 10  
US-07-878-960-1/c  
Sequence 1, Application US/07878960  
Patent No. 5444164  
GENERAL INFORMATION:  
APPLICANT: Purchio, Anthony F.  
APPLICANT: Skonier, John  
TITLE OF INVENTION: Neuberger, Michael G.  
TITLE OF INVENTION: TGF-BETA INDUCED GENE AND PROTEIN  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bristol-Myers Squibb Company  
STREET: 3005 First Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/878,960  
FILING DATE: 05-MAY-1992  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/833,835  
FILING DATE: 05-FEB-1992  
NAME: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Sorrentino, Joseph M.  
REGISTRATION NUMBER: 32,598  
REFERENCE/DOCKET NUMBER: ON0092-  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206/728-4800  
TELEFAX: 206/727-3601  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2691 base pairs





Patent No. 6478825

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnovers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerlitsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, J. Christopher

APPLICANT: Guiney, Austin L.

APPLICANT: Kljavin, Ivar J.

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas P.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

TITLE OF INVENTION: Acids Encoding the Same

FILE REFERENCE: P2730P1C13

CURRENT APPLICATION NUMBER: US/09/996,243

CURRENT FILING DATE: 2001-11-14

PRIOR APPLICATION NUMBER: 60/049787

PRIOR FILING DATE: 1997-06-16

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/065186

PRIOR FILING DATE: 1997-11-12

PRIOR APPLICATION NUMBER: 60/065111

PRIOR FILING DATE: 1997-11-13

PRIOR APPLICATION NUMBER: 60/066770

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/075945

PRIOR FILING DATE: 1998-02-25

PRIOR APPLICATION NUMBER: 60/078910

PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/083322

PRIOR FILING DATE: 1998-04-28

PRIOR APPLICATION NUMBER: 60/084600

PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/087106

PRIOR FILING DATE: 1998-05-28

PRIOR APPLICATION NUMBER: 60/087607

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087609

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087759

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087827

PRIOR FILING DATE: 1998-06-03

PRIOR APPLICATION NUMBER: 60/088021

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088025

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088026

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088028

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088029

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088030

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088033

PRIOR FILING DATE: 1998-06-04

;; PRIOR APPLICATION NUMBER: 60/088326  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088167  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088202  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088212  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088217  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088655  
;; PRIOR FILING DATE: 1998-06-09  
;; PRIOR APPLICATION NUMBER: 60/088734  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088738  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088742  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088824  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088826  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088861  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088876  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089105  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089440  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089512  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089538  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089598  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089600  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089653  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089801  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089908  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429

;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 1.9%; Score 36; DB 4; Length 2854;  
Best Local Similarity 67.1%; Pred. No. 2;  
Matches 51; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
QY 1776 TTAATAACATGTTTCACCTTAATAAGTAACACTTATAATTTTGGAGATCTCGA 1835  
DB 2778 TTAATGCTTACCTCCCTTTAAATTAATGATTAATAATGCTCTTGAATAAAAAA 2837  
QY 1836 AAAAAAAAAAAAAA 1851  
DB 2838 AAAAAAAAAAAAAA 2853

Search completed: December 22, 2003, 23:17:17  
Job time : 129.256 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 14:49:05 ; Search time 579.108 Seconds  
(without alignments)  
10650.710 Million cell updates/sec

Title: US-09-745-763-35  
Perfect score: 1851  
Sequence: 1 GCGTAGCCGCCGAGCTTAGT.....CTGAAAAAAAAAAAAAAAAA 1851

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapept 1.0

Searched: 2211978 seqs, 1666101734 residues 4423956  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCr\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCrUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1851	100.0	1851	9	US-09-745-763-35
2	1846.2	99.7	1863	11	US-09-984-271-28
3	1145	61.9	1778	10	US-09-917-800A-505
4	1098.4	59.3	1134	11	US-09-984-271-95
5	261.8	14.1	357	10	US-09-833-381-1929
6	247.4	13.4	427	10	US-09-833-381-1930
7	233	12.6	234	10	US-09-880-107-3881
8	135	7.3	546	13	US-10-027-632-192834
9	135	7.3	546	14	US-10-027-632-192834
10	91	4.9	134	10	US-09-783-590-12141
11	60.2	3.3	65	13	US-09-908-975-25931
12	60	3.2	60	13	US-09-908-975-8960
13	45.4	2.5	40862	13	US-10-311-455-2046
14	42	2.3	2200	15	US-10-059-585-5
15	41.8	2.3	13202	13	US-10-311-455-1458

C 16	41.2	2.2	328	13	US-10-076-747-72
C 17	41.2	2.2	664	13	US-09-814-353-6384
C 18	41.2	2.2	664	13	US-09-814-353-12661
C 19	41.2	2.2	7384	13	US-10-311-455-720
C 20	41	2.2	65	13	US-09-908-975-1202
C 21	41	2.2	445	13	US-09-814-353-17408
C 22	40.4	2.2	3673778	13	US-10-312-841-2
C 23	40.2	2.2	397	13	US-09-814-353-17565
C 24	40.2	2.2	8946	13	US-10-311-455-884
C 25	40	2.2	646	8	US-08-809-423A-50
C 26	40	2.2	646	13	US-10-271-344-50
C 27	39.6	2.1	881	8	US-08-781-986A-518
C 28	39.6	2.1	3997	15	US-10-198-846-12639
C 29	39.6	2.1	6200	13	US-10-311-455-2187
C 30	39.4	2.1	484	11	US-09-918-995-30889
C 31	39.4	2.1	1363	15	US-10-232-484-5
C 32	39.4	2.1	2319	14	US-10-044-090-579
C 33	39.4	2.1	3322	13	US-09-814-353-20277
C 34	39.4	2.1	11036	13	US-10-240-453-137
C 35	39.4	2.1	11036	15	US-10-239-676-117
C 36	39	2.1	665	13	US-10-027-632-101927
C 37	39	2.1	665	14	US-10-027-632-101927
C 38	38.8	2.1	487	10	US-09-960-352-7373
C 39	38.8	2.1	811	15	US-10-184-644-414
C 40	38.8	2.1	811	15	US-10-184-634-414
C 41	38.8	2.1	1187	13	US-10-027-632-101130
C 42	38.8	2.1	1187	14	US-10-027-632-101130
C 43	38.8	2.1	7928	13	US-10-311-455-67
C 44	38.8	2.1	3673778	13	US-10-312-841-1
C 45	38.6	2.1	516	13	US-09-814-353-17377

ALIGNMENTS

RESULT 1  
US-09-745-763-35  
Sequence 35, Application US/09745763  
Patent No. US020065394A1  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
McCoy, John M.  
Lavallee, Edward R.  
Collins-Racie, Lisa A.  
Evans, Cheryl  
Werberg, David  
Treacy, Maurice  
Spaulding, Vikki  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM  
NUMBER OF SEQUENCES: 219  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/745,763  
FILING DATE: 18-Jun-2000  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Sprunger, Suzanne A.  
REGISTRATION NUMBER: 41,323  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8284  
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 35:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1851 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 SEQUENCE DESCRIPTION: SEQ ID NO: 35:  
 US-09-745-763-35

Query Match 100.0%; Score 1851; DB 9; Length 1851;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1851; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGCATGCGCGGAGCTTACCTGCTGGAGCCGCTCCGTGCGCCGCTGACAGAGCCGCTTA 60
DB 1 GGCATGCGCGGAGCTTACCTGCTGGAGCCGCTCCGTGCGCCGCTGACAGAGCCGCTTA 60
QY 61 TCAGATTATCTTACAGAAAACCAACTGGAATAAATAATGAAATTCCTTATCTTCGAT 120
DB 61 TCAGATTATCTTACAGAAAACCAACTGGAATAAATAATGAAATTCCTTATCTTCGAT 120
QY 121 TTTTCGGTGTGCTTACCTTTTATCTGTGCTGAGAAAGCTATATGCAAGATGCA 180
DB 121 TTTTCGGTGTGCTTACCTTTTATCTGTGCTGAGAAAGCTATATGCAAGATGCA 180
QY 181 TCTTAAGAGACCTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTTGCTA 240
DB 181 TCTTAAGAGACCTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTTGCTA 240
QY 241 AAGCAATGATCAACCTAGCTGTTTATGTTAAGCCAGAACAGATCCTATGAGGATTTG 300
DB 241 AAGCAATGATCAACCTAGCTGTTTATGTTAAGCCAGAACAGATCCTATGAGGATTTG 300
QY 301 CACTTCGTGTTGATCTGTGTGAGCCAGACTGAGTGTCTCAAGAACCTAGAAAAAGCA 360
DB 301 CACTTCGTGTTGATCTGTGTGAGCCAGACTGAGTGTCTCAAGAACCTAGAAAAAGCA 360
QY 361 TCCAATTATGTTACAAAACCTGCGAGCAAGATGGGCTGAGAAAGTTCACTGGAAGCAG 420
DB 361 TCCAATTATGTTACAAAACCTGCGAGCAAGATGGGCTGAGAAAGTTCACTGGAAGCAG 420
QY 421 TGAGAAATCCCACTGGGAGAGGAGAGAAATGACGTGTGATGCTGAGACCAAGATTC 480
DB 421 TGAGAAATCCCACTGGGAGAGGAGAGAAATGACGTGTGATGCTGAGACCAAGATTC 480
QY 481 ATAGATAGCCATCTCGGCTCTTGTGAGAGCAATTTGGAATCTCTCAGAAAGCATTAAG 540
DB 481 ATAGATAGCCATCTCGGCTCTTGTGAGAGCAATTTGGAATCTCTCAGAAAGCATTAAG 540
QY 541 CAGAAATCTGTGTGTGACCTTTTGTGATGAACCTGACAGAAAGGGCTCAGAAAGCAG 600
DB 541 CAGAAATCTGTGTGTGACCTTTTGTGATGAACCTGACAGAAAGGGCTCAGAAAGCAG 600
QY 601 GGAATATGTTGTTTATTAACCAACCTTACATCACTCAAGAGAGGTCATATCCGA 660
DB 601 GGAATATGTTGTTTATTAACCAACCTTACATCACTCAAGAGAGGTCATATCCGA 660
QY 661 CGCAGAGGGGCGGTGAGAGCTGCAAGAGTGGGGCTTTGTGATCTCTCATTCGATCCGTG 720
DB 661 CGCAGAGGGGCGGTGAGAGCTGCAAGAGTGGGGCTTTGTGATCTCTCATTCGATCCGTG 720
QY 721 CCTCTCTTCCATCTACAGTCTCTCAACAGTATTCAGGAATACCAAGATGCGTCCCA 780
DB 721 CCTCTCTTCCATCTACAGTCTCTCAACAGTATTCAGGAATACCAAGATGCGTCCCA 780
QY 781 AAATTCGAACGCGTATTAAGAGTGAAGATGCAAGAAATGATGCAAGAAATGCGTCTC 840
DB 781 AAATTCGAACGCGTATTAAGAGTGAAGATGCAAGAAATGATGCAAGAAATGCGTCTC 840
QY 841 ATGGGATCAAAATTTGTCACTTAAGATGAGGAGCAAAAGACTACCAAGATCTGAT 900
DB 841 ATGGGATCAAAATTTGTCACTTAAGATGAGGAGCAAAAGACTACCAAGATCTGAT 900

```

```

QY 901 CCTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGTTGTACTGTGCA 960
DB 901 CCTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGTTGTACTGTGCA 960
QY 961 GTGACATCTGACAGCTGGAGATGTTGGCAGAGGTGCAATGATGATGCGGTGAGCT 1020
DB 961 GTGACATCTGACAGCTGGAGATGTTGGCAGAGGTGCAATGATGATGCGGTGAGCT 1020
QY 1021 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTTTGGCTGCGTCCAAAGGACTC 1080
DB 1021 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTTTGGCTGCGTCCAAAGGACTC 1080
QY 1081 TGGGCTGTGCTGTGACAGTCAAGAAACAAAGGTGAGTTGAGCTTCAGATTAATC 1140
DB 1081 TGGGCTGTGCTGTGACAGTCAAGAAACAAAGGTGAGTTGAGCTTCAGATTAATC 1140
QY 1141 AGTTACACAAGGTAAATATTTTCAACTACAGTCTGTGATGAGTCTGACGCAAGAACT 1200
DB 1141 AGTTACACAAGGTAAATATTTTCAACTACAGTCTGTGATGAGTCTGACGCAAGAACT 1200
QY 1201 TCTTACCACTGGGCTGCAATTTCACTGCACTGAGTGAAGAGCCAGGCGCATCATGAGAGG 1260
DB 1201 TCTTACCACTGGGCTGCAATTTCACTGCACTGAGTGAAGAGCCAGGCGCATCATGAGAGG 1260
QY 1261 TTATGAGCTGTGCTGACGCTCCCAATATCACTGAGTCTGAGGCTGAGGCAATGAGAGGAG 1320
DB 1261 TTATGAGCTGTGCTGACGCTCCCAATATCACTGAGTCTGAGGCTGAGGCAATGAGAGGAG 1320
QY 1321 ACATCAACTTTTGGATCCAAAGCTGAGTGTGCTGAGGCACTCTGATGATCTTATACA 1380
DB 1321 ACATCAACTTTTGGATCCAAAGCTGAGTGTGCTGAGGCACTCTGATGATCTTATACA 1380
QY 1381 AGTATTTCTTCTTCATCATCTCCCAAGAGACACATGATCTGTATGATCTCAAAAGCAGA 1440
DB 1381 AGTATTTCTTCTTCATCATCTCCCAAGAGACACATGATCTGTATGATCTCAAAAGCAGA 1440
QY 1441 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTTCTTATGTTGTTGCAAGACATGGAAGAA 1500
DB 1441 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTTCTTATGTTGTTGCAAGACATGGAAGAA 1500
QY 1501 TGCTGCTAGGCTCTAGAAAACAGTAAAGAAAAGTATTCATCTTCTGCGCAGAAATC 1560
DB 1501 TGCTGCTAGGCTCTAGAAAACAGTAAAGAAAAGTATTCATCTTCTGCGCAGAAATC 1560
QY 1561 CTGGCTGTGCACTTTGGAAGAACTCTCTTCACTAACAATTTCAATTCATCTTCA 1620
DB 1561 CTGGCTGTGCACTTTGGAAGAACTCTCTTCACTAACAATTTCAATTCATCTTCA 1620
QY 1621 AAGCAACCTATTTTCAAGCTTCTGTATATCTTTCTGATGATCTTCCAAATTCCT 1680
DB 1621 AAGCAACCTATTTTCAAGCTTCTGTATATCTTTCTGATGATCTTCCAAATTCCT 1680
QY 1681 GATTCTAGAAAAAGAAATCATCTTCCCTCCCTCCCAACATAGAAATCAATATGTA 1740
DB 1681 GATTCTAGAAAAAGAAATCATCTTCCCTCCCTCCCAACATAGAAATCAATATGTA 1740
QY 1741 GGAATTAAGTGGGGGCAATTTCTTATATCACTCTTTTAAACATTTGTTTCACTTTAA 1800
DB 1741 GGAATTAAGTGGGGGCAATTTCTTATATCACTCTTTTAAACATTTGTTTCACTTTAA 1800
QY 1801 AGTAAACACTTAATTAATTTTGAAGATCTCTGAGAAAAAATTTTTTTTTT 1851
DB 1801 AGTAAACACTTAATTAATTTTGAAGATCTCTGAGAAAAAATTTTTTTTTT 1851

```

RESULT 2  
 US-09-984-271-28  
 ; Sequence 28, Application US/09984271  
 ; Publication No. US20030040088A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: 71 Human Secreted Proteins

FILE REFERENCE: P2030P1  
 CURRENT APPLICATION NUMBER: US/09/984,271  
 CURRENT FILING DATE: 2001-10-29  
 PRIOR APPLICATION NUMBER: 09/482,273  
 PRIOR FILING DATE: 2000-01-13  
 PRIOR APPLICATION NUMBER: PCT/US99/15849  
 PRIOR FILING DATE: 1999-07-14  
 PRIOR APPLICATION NUMBER: 60/092,921  
 PRIOR FILING DATE: 1998-07-15  
 PRIOR APPLICATION NUMBER: 60/092,922  
 PRIOR FILING DATE: 1998-07-15  
 PRIOR APPLICATION NUMBER: 60/092,956  
 PRIOR FILING DATE: 1998-07-15  
 NUMBER OF SEQ ID NOS: 267  
 SOFTWARE: Patentin Ver. 2.0  
 SEQ ID NO 28  
 LENGTH: 1863  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-09-984-271-28

Query Match 99.7%; Score 1846.2; DB 11; Length 1863;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1848; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 GGCTAGGCGGAGGCTTAGTCTGGAGCCGCTCCGCGCGCGGCGGAGCCGCTTA 60  
 1 GACTAGGCGGAGGCTTAGTCTGGAGCCGCTCCGCGCGCGGCGGAGCCGCTTA 60  
 61 TCAGATTATCTTAAACAAGAAACCACTGAAAGAAATTCCTTATCTTCGAT 120  
 61 TCAGATTATCTTAAACAAGAAACCACTGAAAGAAATTCCTTATCTTCGAT 120  
 121 TTTTGGGTGGTTCACCTTTTATCCTGTGCTCTGSAAGAACTATATGCAAGATGCA 180  
 121 TTTTGGGTGGTTCACCTTTTATCCTGTGCTCTGSAAGAACTATATGCAAGATGCA 180  
 181 TCTCTAAGAGGACTTTTGAAGAAATTAAGAAATGAGCAAGCTGAGAGATGTTGCTA 240  
 181 TCTCTAAGAGGACTTTTGAAGAAATTAAGAAATGAGCAAGCTGAGAGATGTTGCTA 240  
 241 AAGCAATCATCAACTAGCTGTTTATGTTAAAGCCAGAAAGATCTTATGAGCGATTGG 300  
 241 AAGCAATCATCAACTAGCTGTTTATGTTAAAGCCAGAAAGATCTTATGAGCGATTGG 300  
 301 CACTTCTGTGATATCTGTGGAAGCCAGAGTGAAGTCCAGAAAGCTTAAAGAAAGCCA 360  
 301 CACTTCTGTGATATCTGTGGAAGCCAGAGTGAAGTCCAGAAAGCTTAAAGAAAGCCA 360  
 361 TCCAAATATATGATACCAAAACCTGAGCAAGAGTGGCTGGAAGAAATTCACCTGAGCCAG 420  
 361 TCCAAATATATGATACCAAAACCTGAGCAAGAGTGGCTGGAAGAAATTCACCTGAGCCAG 420  
 421 TGAGAAATACCCCACTGAGAGAGGAGAAAGATCACTGTGATGCTGAGAGCCAAATTC 480  
 421 TGAGAAATACCCCACTGAGAGAGGAGAAAGATCACTGTGATGCTGAGAGCCAAATTC 480  
 481 ATAAAGATAGCCATCTGGGTCTTGGCAGAGCATTTGGGACTCTCCAGAAAGCAATTAAG 540  
 481 ATAAAGATAGCCATCTGGGTCTTGGCAGAGCATTTGGGACTCTCCAGAAAGCAATTAAG 540  
 541 CAGAAGTCTGTGATGATCTTTCGATGAACTGCAAGAGAGGAGCTCAGAAAGCAAG 600  
 541 CAGAAGTCTGTGATGATCTTTCGATGAACTGCAAGAGAGGAGCTCAGAAAGCAAG 600  
 601 GGAAGATTTGTTTATTAACCAACTTACATCACTCAAGAGAGGAGTGAATCCGA 660  
 601 GGAAGATTTGTTTATTAACCAACTTACATCACTCAAGAGAGGAGTGAATCCGA 660  
 661 CGCAGAGGAGGAGTGAAGCTGCAAGAGTGGGAGCTTTGGCATCTCTCATTTGATCCGTTGG 720  
 661 CGCAGAGGAGGAGTGAAGCTGCAAGAGTGGGAGCTTTGGCATCTCTCATTTGATCCGTTGG 720

721 CCTCCTTCTCCATCTAGATCTCTACACAGGATTTAGAGAAATCCAGAGATGCGTCCCA 780  
 721 CCTCCTTCTCCATCTAGATCTCTACACAGGATTTAGAGAAATCCAGAGATGCGTCCCA 780  
 781 AATTCACAGAGCTGTATTAAGGTGAAGATGCAAGAAATGATTCAGAAATGCGTTCTC 840  
 781 AATTCACAGAGCTGTATTAAGGTGAAGATGCAAGAAATGATTCAGAAATGCGTTCTC 840  
 841 ATGGGATCAAAATTTGATTCAGTAAAGATGGGGGCAAGAGCTTACCAGATCTGATTT 900  
 841 ATGGGATCAAAATTTGATTCAGTAAAGATGGGGGCAAGAGCTTACCAGATCTGATTT 900  
 901 CCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTTGTACTGTGCA 960  
 901 CCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTTGTACTGTGCA 960  
 961 GTGACATCTGACAGCTGGAGATGTTGGCAGAGGCTGCATGATGATGCGGTGAGCCT 1020  
 961 GTGACATCTGACAGCTGGAGATGTTGGCAGAGGCTGCATGATGATGCGGTGAGCCT 1020  
 1021 TTATATCATGGAGAGCACTCTCACTTATTAAGATCTTGGGCTGCGCCCAAGAGGAGCTC 1080  
 1021 TTATATCATGGAGAGCACTCTCACTTATTAAGATCTTGGGCTGCGCCCAAGAGGAGCTC 1080  
 1081 TGCGGCTGTGCTGTGACTGCAAGAAACAGAGTGAAGTTGGCTTCCAGATTAATC 1140  
 1081 TGCGGCTGTGCTGTGACTGCAAGAAACAGAGTGAAGTTGGCTTCCAGATTAATC 1140  
 1141 AGTTACCAAGGTAATATTTCCAACTACAGCTGTGATGAGTCTGACGAGAACT 1200  
 1141 AGTTACCAAGGTAATATTTCCAACTACAGCTGTGATGAGTCTGACGAGAACT 1200  
 1201 TCTTACCACTGCGCTGCAATTCACCTGCAAGTGAAGAGCCAGCCATCATGAGAGAG 1260  
 1201 TCTTACCACTGCGCTGCAATTCACCTGCAAGTGAAGAGCCAGCCATCATGAGAGAG 1260  
 1261 TTATGAGCTGTGCTGAGCCCTCAATATCACTGAGTCTGAGGAGCATGAGAGAGGAG 1320  
 1261 TTATGAGCTGTGCTGAGCCCTCAATATCACTGAGTCTGAGGAGCATGAGAGAGGAG 1320  
 1321 ACATCAACTTTTGGATCCAACTGAGAGTGGCTGAGGAGCATCTTATGATGATTAATCA 1380  
 1321 ACATCAACTTTTGGATCCAACTGAGAGTGGCTGAGGAGCATCTTATGATGATTAATCA 1380  
 1381 AGTATTTCTTTTCAATCACTCCCAAGAGACCAATGATGATGATCAAAAGCAGA 1440  
 1381 AGTATTTCTTTTCAATCACTCCCAAGAGACCAATGATGATGATCAAAAGCAGA 1440  
 1441 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTAATGTTGTCAGACATGAGAGAA 1500  
 1441 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTAATGTTGTCAGACATGAGAGAA 1500  
 1501 TGCTGCTAGGCTCTAGAAACAGTAAGAAAGAAAGCTTTTATGATCTTCTGCGCAGAAATC 1560  
 1501 TGCTGCTAGGCTCTAGAAACAGTAAGAAAGAAAGCTTTTATGATCTTCTGCGCAGAAATC 1560  
 1561 CTGGGTCTGCAACTTTGAGAAATCTCTTCAATCAATCAATTTATCAATTTCTTCA 1620  
 1561 CTGGGTCTGCAACTTTGAGAAATCTCTTCAATCAATCAATTTATCAATTTCTTCA 1620  
 1621 AAGCAACACTATTTTCAATGCTTCTGTTATTAATCTTCTGATTAATTTCAATTTCT 1680  
 1621 AAGCAACACTATTTTCAATGCTTCTGTTATTAATCTTCTGATTAATTTCAATTTCT 1680  
 1681 GATTCTAGAAAAAGAAATCAATTTCTCCCTCCCTCCCAACATAGAAATCAATATGTTA 1740  
 1681 GATTCTAGAAAAAGAAATCAATTTCTCCCTCCCTCCCAACATAGAAATCAATATGTTA 1740  
 1741 GGAATTAAGTGGGAGCTTTCTTTATATCACTCTTTTAAATCAATTTGTTTCACTTTTAA 1800  
 1741 GGAATTAAGTGGGAGCTTTCTTTATATCACTCTTTTAAATCAATTTGTTTCACTTTTAA 1800  
 1801 AGTAAACACTTAATTAATTTTGGAGAGATCTCTGAAAAAAGAAAAAAGAAAAA 1861



Db 1801 AGTAACTATTAATTTTGGAGAGATCTCGAAAAA 1851

## RESULT 3

US-09-917-800A-505

Sequence 505, Application US/09917800A  
Patent No. US20020119462A1

## GENERAL INFORMATION:

APPLICANT: Mendrick, Donna

APPLICANT: Porter, Mark

APPLICANT: Johnson, Kory

APPLICANT: Castle, Arthur

APPLICANT: Elashoff, Michael

APPLICANT: Gene Logic, Inc.

TITLE OF INVENTION: Molecular Toxicology Modeling

FILE REFERENCE: 44921-5038-US

CURRENT APPLICATION NUMBER: US/09/917, 800A

CURRENT FILING DATE: 2001-07-31

PRIOR APPLICATION NUMBER: US 60/222, 040

PRIOR FILING DATE: 2000-07-31

PRIOR APPLICATION NUMBER: US 60/222, 880

PRIOR FILING DATE: 2000-11-02

PRIOR APPLICATION NUMBER: US 60/290, 029

PRIOR FILING DATE: 2001-05-11

PRIOR APPLICATION NUMBER: US 60/290, 645

PRIOR FILING DATE: 2001-05-15

PRIOR APPLICATION NUMBER: US 60/292, 336

PRIOR FILING DATE: 2001-05-22

PRIOR APPLICATION NUMBER: US 60/295, 798

PRIOR FILING DATE: 2001-06-06

PRIOR APPLICATION NUMBER: US 60/297, 457

PRIOR FILING DATE: 2001-06-13

PRIOR APPLICATION NUMBER: US 60/298, 884

PRIOR FILING DATE: 2001-06-19

PRIOR APPLICATION NUMBER: US 60/303, 459

PRIOR FILING DATE: 2001-07-09

NUMBER OF SEQ ID NOS: 1740

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 505

LENGTH: 1778

TYPE: DNA

ORGANISM: Rattus norvegicus

FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020119462A1 AF097723

US-09-917-800A-505

Query Match 61.9%; Score 1145; DB 10; Length 1778;  
Best Local Similarity 81.9%; Pred. No. 0;  
Matches 1334; Conservative 0; Mismatches 290; Indels 5; Gaps 1;

QY 73 AACAGAAACCACTGAGAAAAAATTCCTATCTTCGATTTTCGGTGGT 132  
Db 95 AGCAAGAAAGAAAGAACTAGGACATAGGTTCTTCTTCCTGTTGCTGTTG 154

QY 133 TTCACCTTTTATCCCTGTCCTGCGAAAGCTATATGCAAGATGCAATCTTAAGAGA 192  
Db 155 TTCACCTTTTCTCCTTGGGCTCTGAGAAAGCTATATGCAAGATGCTTTTCTGCGAA 214

QY 193 CTTTGAAGAAATTAAGAAATTAAGCACTGTCGAGATTTGCTAAAGCAATCATCA 252  
Db 215 CATTTCAGAAATTAAGAAATTAAGCACTGTCGAGATTTGCTAAAGCAATCATCA 274

QY 253 ACCTGCTGTTATGTAAGCCAGAAAGATCCATAGAGATTTGGACCTTCGTTG 312  
Db 275 ACCTGCTGTTATGTAAGAAATTAAGCAAGCCGATGATAGCGTTTGGACCTTCGTTG 334

QY 313 ATACTGTGAGCCAGAGCTGAGTGGCTCAAGAACTTAGAAAGCAATCAATTAAT 372  
Db 335 ATACTGTGAGCCAGAGCTGAGTGGCTCTTAAGAACTTAGAAAGCAATCAATTAAT 394

QY 373 ACCAAACCTGAGCAAGATGAGCTGAGAAAGTTCACTGAGCCAGTGAATATACCC 432  
Db 395

ACTGAGAGAGGAGAGAAATCAGCTGATGATGAGAGCCAAAGATTCATAGATAGCA 492  
ACTGAGAGAGGAGAGAAATCAGCTGATGATGAGAGCCAAAGATTCATAGATAGCA 514  
TCCTGGGCTCTTGGAGAGAGATTTGGAGCTCTCCAGAAAGCATTAAGCAGAGTTCTGG 552  
TTTTAGGCTCTTGGAGAGAGATTTGGAGCTCTCCAGAAAGCATTAAGCAGAGTTCTGG 574  
TGTGACCTCTTGGAGAGAGATTTGGAGCTCTCCAGAAAGCATTAAGCAGAGTTCTGG 612  
TGTGAGCTCTTGGAGAGAGATTTGGAGCTCTCCAGAAAGCATTAAGCAGAGTTCTGG 634  
TTTATTAACCAACTTATCACTACTCAAGAGAGGTCATTAAGCAAGCAGAGGAGGCGG 672  
TTTATTAACCAACTTATCACTACTCAAGAGAGGTCATTAAGCAAGCAGAGGAGGCGG 694  
TGAAGCTGCGAAGGAGGAGGCTTTGGAGCTCTCAATTCAGATTCGAGCTCTTTCTCA 732  
TGAAGCTGCGAAGGAGGAGGCTTTGGAGCTCTCAATTCAGATTCGAGCTCTTTCTCA 754  
TCTACAGCTCTCAACAGATTTTCAAGATTCAGAGATTCGAGCTCTTTCTCAAG 792  
TCTACAGCTCTCAACAGATTTTCAAGATTCAGAGATTCGAGCTCTTTCTCAAG 814  
CCTGATTAACCGTGAAGATTCAGAGATTCAGAGATTCGAGCTCTTTCTCAAG 852  
CCTGATTAACCGTGAAGATTCAGAGATTCAGAGATTCGAGCTCTTTCTCAAG 874  
TTGTATTCATCAATTAAGATTCAGAGATTCAGAGATTCGAGCTCTTTCTCAAG 912  
TTGTATTCATCAATTAAGATTCAGAGATTCAGAGATTCGAGCTCTTTCTCAAG 934  
TAGCAGAGATTCAGAGATTCAGAGATTCAGAGATTCGAGCTCTTTCTCAAG 972  
TTGAGAGATTCAGAGATTCAGAGATTCAGAGATTCGAGCTCTTTCTCAAG 994  
ACAGCTGGAGTGTGGGAGGAGGCTCCAGATTCAGAGATTCGAGCTCTTTCTCAAG 1032  
ACAGCTGGAGTGTGGGAGGAGGCTCCAGATTCAGAGATTCGAGCTCTTTCTCAAG 1054  
AAGCACTCTCACTTAATTAAGATTCGAGCTCTCCAGATTCAGAGATTCGAGCTCTTTCTCAAG 1092  
AAGCACTCTCACTTAATTAAGATTCGAGCTCTCCAGATTCAGAGATTCGAGCTCTTTCTCAAG 1114  
TCTGAGCTGCGAAGATTCAGAGATTCAGAGATTCGAGCTCTTTCTCAAG 1152  
TCTGAGCTGCGAAGATTCAGAGATTCAGAGATTCGAGCTCTTTCTCAAG 1174  
TCTGAGCTGCGAAGATTCAGAGATTCAGAGATTCGAGCTCTTTCTCAAG 1212  
TCTGAGCTGCGAAGATTCAGAGATTCAGAGATTCGAGCTCTTTCTCAAG 1234  
GCTGCAATTCAGCTGAGTGAAGAGGCAAGGCTCATGAGAGAGGTTATGAGCTGCG 1272  
GCTGCAATTCAGCTGAGTGAAGAGGCAAGGCTCATGAGAGAGGTTATGAGCTGCG 1294  
GCTGCAATTCAGCTGAGTGAAGAGGCAAGGCTCATGAGAGAGGTTATGAGCTGCG 1332  
TGCAGCTGCTCAATTAATTCAGAGTCTGAGAGCTCATGAGAGAGGTTATGAGCTGCG 1354  
TGCAGCTGCTCAATTAATTCAGAGTCTGAGAGCTCATGAGAGAGGTTATGAGCTGCG 1392  
GGATTCAGAGTGAAGGCTGAGAGCTCATGAGAGAGGTTATGAGCTGCG 1414  
GGATTCAGAGTGAAGGCTGAGAGCTCATGAGAGAGGTTATGAGCTGCG 1434  
TGCATCACTCTCCAGAGAGCAATGATCTGATGATTCAGAGAGGTTATGAGCTGCG 1452  
TGCATCACTCTCCAGAGAGCAATGATCTGATGATTCAGAGAGGTTATGAGCTGCG 1474  
TGCATCACTCTCCAGAGAGCAATGATCTGATGATTCAGAGAGGTTATGAGCTGCG 1512  
CTGCTGTTGGGCTGTTGTTCTTAATGTTGTTGAGAGATTCAGAGAGGTTATGAGCTGCG 1534  
CTGCTGTTGGGCTGTTGTTCTTAATGTTGTTGAGAGATTCAGAGAGGTTATGAGCTGCG 1554

Oy	1513	CTTGAAGAACTAAGAAAGAAACGTTTTCATGCTCTGGCCAGGAATCTGGGCTGCA	15172
Db	1535	CTTAAAGGAAACAAAGAGAGAAAGAACCTTGTCTCTGCAAGCTGGGATCCCATTTGGGAT	1594
Oy	1573	CTTTGAAAACCTCCTCTTCACATACAAATTTTCATCATTCACTCTTCAAGACACAATCT	1632
Db	1595	TTTCACAGCAGCAATCTTCAAGACAGCTGTATTATACATCAATCCCGCTGGACAC-----	1649
Oy	1633	ATTTCATGCTTTCTGTATTATCTTTCTTGATACCTTCCAAATTCCTGATTTCTAGAAAA	16922
Db	1650	GTTTCTTATTACCTTCTGTAAACATCTTTCCTGATACGCTTTTACCTGTGTTCTAGATA	1709
Oy	1693	AGGATATCAT	1701
Db	1710	AGTATATCAT	1718

## RESULT

```

US-09-984-271-95
Sequence 95, Application US/09984271
Publication No. US20030040088A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/984,271
PRIORITY FILING DATE: 2001-10-29
PRIORITY APPLICATION NUMBER: 09/482,273
PRIORITY FILING DATE: 2000-01-13.
PRIORITY APPLICATION NUMBER: PCT/US99/15849
PRIORITY FILING DATE: 1999-07-14
PRIORITY APPLICATION NUMBER: 60/092,921
PRIORITY FILING DATE: 1998-07-15
PRIORITY APPLICATION NUMBER: 60/092,922
PRIORITY FILING DATE: 1998-07-15
PRIORITY APPLICATION NUMBER: 60/092,956
PRIORITY FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 95
LENGTH: 1134
TYPE: DNA
ORGANISM: Homo sapiens
US-09-984-271-95

```

Query Match	59.3%;	Score 1098.4;	DB 11;	Length 1134;
-------------	--------	---------------	--------	--------------

Best Local Similarity 59.7%; P-Seq. NO. 0;  
Matches 1121; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Oy	729	TCGATCTCAGGCTCTCACACAGGTATTCAGGAAATCACAGATATGGCGTCCCAAAATTCCA	788
Db	1	TCGATCTCAGGCTCTCACACAGGTATTCAGGAAATCACAGATATGGCGTCCCAAAATTCCA	60
Oy	789	ACAGCTGTATTTACGGTGGAAAGATCAGAAATGATGCAAGAAATGGCTTCTTCATGGGATC	848
Db	61	ACAGCTGTATTTACGGTGGAAAGATCAGAAATGATGCAAGAAATGGCTTCTTCATGGGATC	120
Oy	849	AAAAATTCATTCAGCTTAAAGATGGGGCCAAAGACTTACCCAGATATTCGATCTTCTTCAAC	908
Db	121	AAAAATTCATTCAGCTTAAAGATGGGGCCAAAGACTTACCCAGATATTCGATCTTCTTCAAC	180
Oy	909	ACTGTAGCAGAGATCACTGGGAGCAAAATATCCAGAAACAGTTGTACTGGTCAGTGGACAT	968
Db	181	ACTGTAGCAGAGATCACTGGGAGCAAAATATCCAGAAACAGTTGTACTGGTCAGTGGACAT	240
Oy	969	CTGGACAGCTGGATGTTGGGCGAGGATGCATGATGATGGCGGTGAGACCTTTATATCA	1028
Db	241	CTGGACAGCTGGAGTGTGGGCGAGGATGCATGATGATGGCGGTGAGACCTTTATATCA	300
Oy	1029	TGGGAAGCACTTCATTTAAATCTTGGGCGTGCCTCCAAAGAAGACTCTGGGCGGTG	1088
Db	301	TGGGAAGCACTTCATTTAAAGATCTTGGGCGTGCCTCCAAAGAAGACTCTGGGCGGTG	360

OY	1089	GTGCTCTGGA	CTGAGAAAGAA	CAAGGAGAGT	GGTGCTCTCCAGTTATACATTCAC	1148
Db	361	GTGCTCTGGA	CTGAGAAAGAA	CAAGGAGAGT	GGTGCTCTCCAGTTATACATTCAC	420
OY	1149	AAGGTAATAT	ATTTCCAACTA	CAGTCTGTGATG	AGTCTGACGAGAAACCTTCTTACC	1208
Db	421	AAGGTAATAT	ATTTCCAACTA	CAGTCTGTGATG	AGTCTGACGAGAAACCTTCTTACC	480
OY	1209	ACTGGGCTGA	TCATTCATCTGG	CAGTGAAGGCC	CATCATGAGAGGTTATGAC	1268
Db	481	ACTGGGCTGA	TCATTCATCTGG	CAGTGAAGGCC	CATCATGAGAGGTTATGAGC	539
OY	1269	CTGCTGAGCC	CCCTCAATATCA	CTCAAGTCTTG	AGCCATGGAGAAAGGACAGACATCAAC	1328
Db	540	CTGCTGAGCC	CCCTCAATATCA	CTCAAGTCTTG	AGCCATGGAGAAAGGACAGACATCAAC	599
OY	1329	TTTTGATCCA	AGCTGAGAGTGC	CTGAGCCAGTCA	CTTAGTGACTTATCAAGTATTTCC	1388
Db	600	TTTTGATCCA	AGCTGAGAGTGC	CTGAGCCAGTCA	CTTAGTGACTTATCAAGTATTTCC	659
OY	1389	TTCTTCCATC	ACTCTCCACG	AGACACCATGA	CTGTCAATGATCCAAAGCAGATGAATGTT	1448
Db	660	TTCTTCCATC	ACTCTCCACG	AGACACCATGA	CTGTCAATGATCCAAAGCAGATGAATGTT	719
OY	1449	GTGCTGCTGT	TTTGGGCGTGTT	TTCTTATGTTGTG	GACATGGAAGAAATCGTGCCT	1508
Db	720	GTGCTGCTGT	TTTGGGCGTGTT	TTCTTATGTTGTG	GACATGGAAGAAATCGTGCCT	779
OY	1509	AGGTCCTAGA	AAACAGTAAGAA	AGATTTTCATGCTT	CTGGCCAGAGAACTCTGGGTCCT	1568
Db	780	AGGTCCTAGA	AAACAGTAAGAA	AGATTTTCATGCTT	CTGGCCAGAGAACTCTGGGTCCT	839
OY	1569	GCAACTTTGG	AAAACTCCTCTT	TCACATPAACATTT	CAATTCATCTTCAAAGCACA	1628
Db	840	GCAACTTTGG	AAAACTCCTCTT	TCACATPAACATTT	CAATTCATCTTCAAAGCACA	899
OY	1629	CTCATATTTCA	GCCTTTCGTATATAT	CTTTCTTGAATCTT	CCAAATCTCTG-ATTCTA	1687
Db	900	CTCATATTTCA	GCCTTTCGTATATAT	CTTTCTTGAATCTT	CCAAATCTCTCTGCATTTCTA	959
OY	1688	GAAAAAGGAA	TCAATTCCTCCCT	CCCTCCACACATG	AATCAACATATGAGTAGGATTA	1747
Db	960	GAAAAAGGAA	TCAATTCCTCCCT	CCCTCCACACATG	AATCAACATATGAGTAGGATTA	1019
OY	1748	CAGTGGGGGCA	CTTTCTTATATAC	CTCTTAAAAACAT	TTGTTTCACTTTAAAGTAAAC	1807
Db	1020	CAGTGGGGGCA	CTTTCTTATATAC	CTCTTAAAAACAT	TTGTTTCCACTTTAAAGTAAAC	1079
OY	1808	ACTTAATATA	TTTTTGGAGAGTCT	CTGAAAAA	AAAAAAAAAAAAA	1851
Db	1080	ACTTAATATA	TTTTTGGAGAGTCT	CTGAAAAA	AAAAAAAAAAAAA	1123

## RESULT 5

```

US-09-833-381-1929
; Sequence 1929, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO. 1929
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```





Sequence 5, Application US/10059297A1, Publication No. US20030082776A1

GENERAL INFORMATION:

APPLICANT: Ota, Toshio

APPLICANT: Isegai, Taka

APPLICANT: Nishikawa, Tetsuo

APPLICANT: Hayashi, Koji

APPLICANT: Otsuka, Koaru

APPLICANT: Yamamoto, Jun-ichi

APPLICANT: Iehi, Shizuko

APPLICANT: Sugiyama, Tomoya

APPLICANT: Wakamatsu, Ai

APPLICANT: Nagai, Keiichi

APPLICANT: Otsuki, Tetsuji

APPLICANT: Furahashi, Shin-ichi

APPLICANT: Senoo, Chikaki

APPLICANT: Nezu, Jun-ichi

; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN  
 ; TITLE OF INVENTION: KINASE/PROTEIN PHOSPHATASE  
 ; FILE REFERENCE: 06501-098001  
 ; CURRENT APPLICATION NUMBER: US/10/059,585  
 ; CURRENT FILING DATE: 2002-01-29  
 ; PRIOR APPLICATION NUMBER: PCT/JP00/05060  
 ; PRIOR FILING DATE: 2000-07-28  
 ; PRIOR APPLICATION NUMBER: US 60/183,322  
 ; PRIOR FILING DATE: 2000-02-17  
 ; PRIOR APPLICATION NUMBER: US 60/159,590  
 ; PRIOR FILING DATE: 1999-10-18  
 ; PRIOR APPLICATION NUMBER: JP 2000-118776  
 ; PRIOR FILING DATE: 2000-01-11  
 ; PRIOR APPLICATION NUMBER: JP 2000-183767  
 ; PRIOR FILING DATE: 2000-05-02  
 ; PRIOR APPLICATION NUMBER: JP 11-248036  
 ; PRIOR FILING DATE: 1999-07-29  
 ; NUMBER OF SEQ ID NOS: 64  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 5  
 ; LENGTH: 2200  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; NAME/KEY: CDS  
 ; LOCATION: (173)...(1450)  
 ; US-10-059-585-5

Query Match 2.3%; Score 42; DB 15; Length 2200;  
 Best Local Similarity 45.5%; Pred. No. 0.49;  
 Matches 150; Conservative 0; Mismatches 180; Indels 0; Gaps 0;

QY 933 AAATATCCAGAACAGGTTGTACTGTGTCAGTACATCGACAGCTGGAGTTGGGAG 992  
 DB 158 AAGGATGAGGAGGAGTGTGTCAGAGGAAAGCTGCTGATTTTCATGAGGAG 217  
 QY 993 GGTGCATGATGATGAGGCTGAGCCCTTTATATCATGGAAGCACTCTCACTATTAA 1052  
 DB 218 GCAGCCCTGACCGACGCGAGTGGGCGGGAGGAGGAGGCTGCTCTTCAAGG 277  
 QY 1053 GATCTTGGGCTGCGTCCAAAGAGACTGTGCGGCTGCTGTGACTGCAGAAACAA 1112  
 DB 278 GAGAGAGGGCGACCTCACTGTCAGAGTGTGCTGCTCTCGAATGAAAGAGAA 337  
 QY 1113 GGTGAGTGTGCTGCTTCACTATATATGATTAACAAGTAATTTCCAACTACAGT 1172  
 DB 338 GAGAGAGGAGTGTGCTGAGTCTTTGTACCAACCAAGCCCTCTGCGCCGACTCAGC 397  
 QY 1173 CTGGTATGAGTGTGACGACGAACCTTTTACCCTGAGGCTGCAATTCATGAGCAGT 1232  
 DB 398 ATCCCTGCTCTTATATACAGAGCTCCGACAAACACAGCCCTGAGATGCTGACCG 457  
 QY 1233 GAAAAGGCCAGGCGCCATCATGAGAGGTT 1262  
 DB 458 GAGAAACAGTTTGTGTAAAGTGAAGT 487

RESULT 15  
 US-10-311-455-1458/c  
 ; Sequence 1458, Application US/10311455  
 ; Publication No. US20030143606a1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OLEK, Alexander  
 ; APPLICANT: PIEPENBROCK, Christian  
 ; APPLICANT: BERLIN, Kurt  
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation  
 ; FILE REFERENCE: 5013.1014  
 ; CURRENT APPLICATION NUMBER: US/10/311,455  
 ; PRIOR FILING DATE: 2002-12-16  
 ; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
 ; PRIOR FILING DATE: 2001-07-02  
 ; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: DE 10043826.1  
 ; PRIOR FILING DATE: 2000-09-01  
 ; NUMBER OF SEQ ID NOS: 2424  
 ; SEQ ID NO 1458  
 ; LENGTH: 13202  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 ; US-10-311-455-1458

Query Match 2.3%; Score 41.8; DB 13; Length 13202;  
 Best Local Similarity 49.3%; Pred. No. 2;  
 Matches 109; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 1631 CTATTCATGCTTTCGTATATATCTTTGATATCTTCCAAATTCCTGATTCTAGAA 1690  
 DB 3058 CCAATTCCTCCATCTTAAATTTCTTAAATACCACTTAAATTAACCAATTCATATA 2999  
 QY 1691 AAAGATATCTTCCCTCCCTCCCAACATAGATCAATATGATGAGTATACAG 1750  
 DB 2998 AATCTAAATATCTTCTTCTTACCTCCATCTTAAATATCTGAAATAAATATAAT 2939  
 QY 1751 TGGGGGATTTCTTATATACCTCTTAAACAAATGTTTCACTTTAAAGTAAACACT 1810  
 DB 2938 TATTAATAACCCGAAACTTACCTACATACATCTTACCTTCAACATCTACAAACC 2879  
 QY 1811 TAATTAATTTTGGAGATCTTGAAAAAATTTAAAAAATTTAAAAAATTTAAAAA 1851  
 DB 2878 TAATATAACTCTAAAAAATTTATATATATTAATTAATTAATTAATTAATTAAT 2838

Search completed: December 22, 2003, 23:44:36  
 Job time : 589.358 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:18:08 ; Search time 511.6 Seconds

(without alignments)  
11332.658 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_99\_1514

Perfect score: 1416

Sequence: 1 ATGAATTCCTTATCTTCTGC.....AAGAAAGCTGCTAGCTCC 1416

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl1.\*  
1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_hcg\_hum.\*  
31: em\_hcg\_inv.\*  
32: em\_hcg\_other.\*  
33: em\_hcg\_mus.\*  
34: em\_hcg\_pln.\*  
35: em\_hcg\_rtd.\*  
36: em\_hcg\_mam.\*  
37: em\_hcg\_vrt.\*  
38: em\_hcg\_hum.\*  
39: em\_hcg\_mus.\*  
40: em\_hcg\_rtd.\*  
41: em\_hcg\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1416	100.0	1851	6	BD106411 Secreted
2	1412.8	99.8	1778	6	BD105816 Secretory
3	1412.8	99.8	1860	6	AX006440 Sequence
4	1412.8	99.8	1928	9	BC020689 Homo sapi
5	1409.6	99.5	1472	9	AF107834 Homo sapi
6	1408	99.4	1860	9	BD127520 Primer fo
7	1408	99.4	1860	9	AK075132 Homo sapi
8	1399.2	98.8	1794	9	AF119386 Homo sapi
9	1116.8	78.9	1796	10	BC037067 Mus muscu
10	1115.2	78.8	1596	10	AF107835 Mus muscu
11	1107.2	78.2	1726	10	AF110777 Rattus no
12	1107.2	78.2	1778	6	AX400829 Sequence
13	1107.2	78.2	1778	10	AF097723 Rattus no
14	1030.8	72.8	1716	10	AF009513 Mus muscu
15	520.6	36.8	609	6	BD059610 Secreted
16	470.2	33.2	642	6	BD125210 Primer fo
17	470.2	33.2	642	6	BD126476 Primer fo
18	436.4	30.8	15927	9	AC006278 Homo sapi
19	436.4	30.8	15927	9	AP006278 Homo sapi
20	436.4	30.8	171301	9	AC010859 Homo sapi
21	421.4	29.8	423	9	AF107833 Homo sapi
22	331	23.4	417	10	AF107832 Mus muscu
23	325.8	23.0	167883	2	AC121026 Rattus no
24	325.8	23.0	218874	2	AC110103 Rattus no
25	323.2	22.8	214287	2	AC128951 Mus muscu
26	284.8	20.1	314	6	BD058492 Secreted
27	276.6	19.5	462	6	BD105811 Secretory
28	276.6	19.5	462	6	E62849 Secretory P
29	212.6	15.0	317	6	BD076934 5' EST of
30	212.2	15.0	144703	9	AP003112 Homo sapi
31	205	14.5	152176	9	AP003117 Homo sapi
32	205	14.5	171035	9	AP003111 Homo sapi
33	174.4	12.3	214287	2	AC128951 Mus muscu
34	172.8	12.2	203469	2	AC123654 Mus muscu
35	171.2	12.1	218874	2	AC110103 Rattus no
36	163.4	11.5	166050	2	AC013817 Homo sapi
37	163.4	11.5	177008	2	AC011175 Homo sapi
38	163.4	11.5	240178	2	AC094173 Rattus no
39	163.4	11.5	288136	2	AC120489 Rattus no
40	154.4	10.9	238034	2	AC123681 Mus muscu
41	116	8.2	10792	1	AE015790 Shewanell
42	115	8.1	110000	2	AL390202_01 Continuation (2 of
43	114.2	8.1	492	6	BD125961 Primer fo
44	99.2	7.0	10487	1	AE005922 Caulobact
45	96.4	6.8	1729	3	AF077194 Acanthoch

## ALIGNMENTS

RESULT 1  
BD106411 1851 bp DNA linear PAT 18-SEP-2002  
LOCUS BD106411  
DEFINITION Secreted proteins and polynucleotides encoding them.  
ACCESSION BD106411  
VERSION BD106411.1 GI:23201229  
KEYWORDS JP 2002503955-A/2.  
SOURCE Chlamydia sp.  
ORGANISM Chlamydia sp.  
REFERENCE 1 (bases 1 to 1851)  
AUTHORS Jacobs,K., McCoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D.,  
Treacy,M., Spaulding,V., and Agostino,M.J.  
TITLE Secreted proteins and polynucleotides encoding them  
JOURNAL JP 2002503955-A 2 05-FEB-2002;

COMMENT

GENETICS INSTITUTE INC  
 PN JP 2002503955-A/2  
 PD 05-FEB-2002  
 PR 20-MAR-1998 JP 1998545874  
 PR 21-MAR-1997 US 08/822167, 19-MAR-1998 US 09/044466 PI  
 KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI  
 DAVID MERBERG,  
 PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC  
 C12N15/12, C07K14/47, A61K38/17  
 CC Strandedness: Double;  
 CC Topology: Linear;  
 FH Key Location/Qualifiers.

FEATURES  
 source 1..1851  
 /organism="Chlamydia sp."  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:35827"

BASE COUNT 531 a 413 c 438 g 469 t

ORIGIN

Query Match 100.0%; Score 1416; DB 6; Length 1851;  
 Beat Local Similarity 100.0%; Pred. No. 0;  
 Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAATTCCTTATCTTGCATTTTTCGGTGGTTTCACTTTTATCCCTGCTCTGGG 60  
 99 ATGAATTCCTTATCTTGCATTTTTCGGTGGTTTCACTTTTATCCCTGCTCTGGG 158  
 61 AAAGCTATATGCAAGATGCGATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATA 120  
 159 AAAGCTATATGCAAGATGCGATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATA 218  
 121 GCCAGCTGTGAGATGTTGCTTAAGACATCATCACTGCTTTATGTTAAGAGCCAG 180  
 219 GCCAGCTGTGAGATGTTGCTTAAGACATCATCACTGCTTTATGTTAAGAGCCAG 278  
 181 AACAGATCTTATGAGCGATTTGGCACTTCTGTTGATCTGTTGAGACCGAGCTGAGTGC 240  
 279 AACAGATCTTATGAGCGATTTGGCACTTCTGTTGATCTGTTGAGACCGAGCTGAGTGC 338  
 241 TCCAAGAACTTAAGAAAGCCATCCAAATTAATGATGATCAAAACCTGACAGAAAGTGGCTG 300  
 339 TCCAAGAACTTAAGAAAGCCATCCAAATTAATGATGATGATCAAAACCTGACAGAAAGTGGCTG 398  
 301 GAGAAAGTTCACTGAGACCGAGTGAATATCCCACTGAGAGAGGAGAGAAATCACT 360  
 399 GAGAAAGTTCACTGAGACCGAGTGAATATCCCACTGAGAGAGGAGAGAAATCACT 458  
 361 GTGATGTGAGACCGAAATTCATTAAGATAGCCATCTGAGTCTTGGCAGAGCAATTGGG 420  
 459 GTGATGTGAGACCGAAATTCATTAAGATAGCCATCTGAGTCTTGGCAGAGCAATTGGG 518  
 421 ACTCTCTCAAGAGGATTAACAGAGATTTCTGTTGATCTCTTTGATGAATCTGAG 480  
 519 ACTCTCTCAAGAGGATTAACAGAGATTTCTGTTGATCTCTTTGATGAATCTGAG 578  
 481 AGAAGGGCTTCAAGAGGAGGAGAGATGTTGTTTATACCAACTTATCATCACTAC 540  
 579 AGAAGGGCTTCAAGAGGAGGAGAGATGTTGTTTATACCAACTTATCATCACTAC 638  
 541 TCAAGAGCGGTGCAATACCGAAGCGAGGGGCGGTGAAGCTGCCAAGTGGGGCTTTG 600  
 639 TCAAGAGCGGTGCAATACCGAAGCGAGGGGCGGTGAAGCTGCCAAGTGGGGCTTTG 698  
 601 GCATCTCTCATTCGATTCGCTGCTCTCTTCTCATCTCAAGTCTTCAACAGATATTCAG 660  
 699 GCATCTCTCATTCGATTCGCTGCTCTCTTCTCATCTCAAGTCTTCAACAGATATTCAG 758  
 661 GAATATCCAGATGCGGTGCGCCAAATTTCAACAGCTGTATTTACGTTGAGAGATCAAGAA 720  
 759 GAATATCCAGATGCGGTGCGCCAAATTTCAACAGCTGTATTTACGTTGAGAGATCAAGAA 818  
 721 ATGATGTCAAGAAATGGCTTCTCATGAGATCAAAATTTGATTCACGCTAAGATGGGGCA 780

819 ATGATGTCAAGAAATGGCTTCTCATGAGATCAAAATTTCTCATTCAGTAAAGATGGGGCA 878  
 781 AAGACCTACCCAGATATCTGATTTCTTCAACACTGTAGCAGAGATCATCTGGAGCAATAT 840  
 879 AAGACCTACCCAGATATCTGATTTCTTCAACACTGTAGCAGAGATCATCTGGAGCAATAT 938  
 841 CCAGAACAGGTTGATCTGATGATGACATCTGAGACAGCTGGAGATGTTGGGACAGGTGCC 900  
 939 CCAGAACAGGTTGATCTGATGATGACATCTGAGACAGCTGGAGATGTTGGGACAGGTGCC 998  
 901 ATGATGATGAGCGGTGAGGCTTTTATATCATGAGAGCACTTCTCATTTATTAAGATCTT 960  
 999 ATGATGATGAGCGGTGAGGCTTTTATATCATGAGAGCACTTCTCATTTATTAAGATCTT 1058  
 961 GGGCTGCTCCAAAGAGACTCTGCGGCTGTGTCTGAGCTGAGAGAGAAAGAGTGA 1020  
 1059 GGGCTGCTCCAAAGAGACTCTGCGGCTGTGTCTGAGCTGAGAGAGAAAGAGTGA 1118  
 1021 GTTGTGCTCTCCAGATATATCATGATTAACAGAGTAATATTTCCACTACAGTCTGTG 1080  
 1119 GTTGTGCTCTCCAGATATATCATGATTAACAGAGTAATATTTCCACTACAGTCTGTG 1178  
 1081 ATGAGATCTGACGAGAACCTTTTACCCACTGAGCTGCAATTCAGTGCAGTGAAG 1140  
 1179 ATGAGATCTGACGAGAACCTTTTACCCACTGAGCTGCAATTCAGTGCAGTGAAG 1238  
 1141 GCCAGGGCCATCATGAGAGAGGTTATAGCTGTGAGAGCCCTCATATATCATCACTAGTC 1200  
 1239 GCCAGGGCCATCATGAGAGAGGTTATAGCTGTGAGAGCCCTCATATATCATCACTAGTC 1298  
 1201 CTGAGCCATGAGAGAGAGGAGACAGATCATCACTTTTGAATCCAGGTGAGTGCCTGAGACC 1260  
 1299 CTGAGCCATGAGAGAGGAGACAGATCATCACTTTTGAATCCAGGTGAGTGCCTGAGACC 1358  
 1261 AGTCTACTGATGATCAATTAACAGATATTTCTTTCATCATCACTGCCAGAGACCATG 1320  
 1359 AGTCTACTGATGATCAATTAACAGATATTTCTTTCATCATCACTGCCAGAGACCATG 1418  
 1321 ACTGTCTGATGATCAATTAACAGATATTTCTTTCATCATCACTGCCAGAGACCATG 1380  
 1419 ACTGTCTGATGATCAATTAACAGATATTTCTTTCATCATCACTGCCAGAGACCATG 1478  
 1381 GTTGTGAGACAGTGAAGAAATGCTGCTAGTGC 1416  
 1479 GTTGTGAGACAGTGAAGAAATGCTGCTAGTGC 1514

RESULT 2  
 BD105816  
 LOCUS  
 DEFINITION  
 SECRETORY PROTEIN.  
 ACCESSION  
 BD105816.1 GI:23200634  
 VERSION  
 JP 2002502234-A/7.  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 SYNTHETIC CONSTRUCT  
 artificial sequences.  
 1 (bases 1 to 1778)  
 JACOBS, K., MCCOY, J. M., RACIE, L. A., LAVALLIE, E. R., MERBERG, D. and  
 SPAULDING, V.  
 Secretory protein.  
 Patent: JP 2002502234-A 7 22-JAN-2002;  
 TITLE  
 JOURNAL  
 COMMENT

GENETICS INSTITUTE INC  
 PN JP 2002502234-A/7  
 PD 22-JAN-2002  
 PR 16-APR-1997 JP 1997537384  
 PR 18-APR-1996 US 08/634325, 13-JAN-1997 US 08/783520 PI  
 KENNETH JACOBS, JOHN M MCCOY, LISA A RACIE, EDWARD R LAVALLIE, PI  
 DAVID MERBERG,  
 PI VIKKI SPAULDING  
 PC C12N15/12, C07K14/47, A61K38/17  
 CC Strandedness: Double;

CC Topology: Linear; Location/Qualifiers.  
 FH Key Location/Qualifiers  
 FEATURES  
 source 1..1778  
 /organism="synthetic construct"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32630"  
 BASE COUNT 514 a 386 c 422 g 456 t  
 ORIGIN

Query Match 99.8%; Score 1412.8; DB 6; Length 1778;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 ATGAAATTCCTTATTCGATTTTGGTGGTGTGACCTTTATCCCTGCTGGG 60
Db 17 ATGAAATTCCTTATTCGATTTTGGTGGTGTGACCTTTATCCCTGCTGGG 76
Qy 61 AAAGCTATATGCAAGAAATGGCATCTTAAGAGACTTTTGAAGAAATTAAGAATA 120
Db 77 AAAGCTATATGCAAGAAATGGCATCTTAAGAGACTTTTGAAGAAATTAAGAATA 136
Qy 121 GCCAGCTGTGAGATGTTGCTAAAGCAATCAACTAGCTGTTTATGTAAGCCAG 180
Db 137 GCCAGCTGTGAGATGTTGCTAAAGCAATCAACTAGCTGTTTATGTAAGCCAG 196
Qy 181 AACGATCTATGAGCGATGGCACTTGTGTATCTGTGTGACCCAGCTAGTGGC 240
Db 197 AACGATCTATGAGCGATGGCACTTGTGTATCTGTGTGACCCAGCTAGTGGC 256
Qy 241 TCCAAAGACTAGAAAAAGCCATCAATTAATGACCAAACTGACAGATGGGCTG 300
Db 257 TCCAAAGACTAGAAAAAGCCATCAATTAATGACCAAACTGACAGATGGGCTG 316
Qy 301 GAGAAAGTTCACTGAGGCCAGTGAATAATCCCACTGGAGAGAGAAATCAAGCT 360
Db 317 GAGAAAGTTCACTGAGGCCAGTGAATAATCCCACTGGAGAGAGAAATCAAGCT 376
Qy 361 GTGATGCTGAGCGCAAGATTCATAGATGACCATCTGGGTCTTGGACAGCATGGG 420
Db 377 GTGATGCTGAGCGCAAGATTCATAGATGACCATCTGGGTCTTGGACAGCATGGG 436
Qy 421 ACTCTCTCAGAAAGCAATTACAGAGAAATTTGTTGAGACCTCTTTCAGATGAG 480
Db 437 ACTCTCTCAGAAAGCAATTACAGAGAAATTTGTTGAGACCTCTTTCAGATGAG 496
Qy 481 AGAAGGCTCTCAGAAAGAGAGAGATGTTGTTATTAACAACCTTACATCACTAC 540
Db 497 AGAAGGCTCTCAGAAAGAGAGAGATGTTGTTATTAACAACCTTACATCACTAC 556
Qy 541 TCAAGAGCGGTGCAATACCAAGCAGAGGGGGGTGAAAGTGGCAAGGTGGGGCTTTG 600
Db 557 TCAAGAGCGGTGCAATACCAAGCAGAGGGGGGTGAAAGTGGCAAGGTGGGGCTTTG 616
Qy 601 GCATCTCTCATTCGATCGTGGCTCTCTTCTCATCTCACTACCTCAACAGATTCAG 660
Db 617 GCATCTCTCATTCGATCGTGGCTCTCTTCTCATCTCACTACCTCAACAGATTCAG 676
Qy 661 GAATACAGAGATGGCGCCCAAAATTCACAAGCTGTATTAAGTGAAGATGACGAA 720
Db 677 GAATACAGAGATGGCGCCCAAAATTCACAAGCTGTATTAAGTGAAGATGACGAA 736
Qy 721 ATGATGTCAGAAATGGCTTCTCATGAGATCAAAATTTGTCATTCAGCTTAAGATGGGGCA 780
Db 737 ATGATGTCAGAAATGGCTTCTCATGAGATCAAAATTTGTCATTCAGCTTAAGATGGGGCA 796
Qy 781 AAAGCTACCAAGATGATGATTCCTTCAACACTGTAGAGAGATCACTGGGAGCAATAT 840
Db 797 AAAGCTACCAAGATGATGATTCCTTCAACACTGTAGAGAGATCACTGGGAGCAATAT 856
Qy 841 CCAGAACAGGTTGATGCTGATGAGCATCTGAGACAGCTGGAGATGTTGGCAGGGTGGC 900
Db 857 CCAGAACAGGTTGATGCTGATGAGCATCTGAGACAGCTGGAGATGTTGGCAGGGTGGC 916

```

Qy 901 ATGATGATGCGCGGTGAGGCTTTATATCATGGAGACACTCTGACTTATTAAGATCTT 960  
 Db 917 ATGATGATGCGCGGTGAGGCTTTATATCATGGAGACACTCTGACTTATTAAGATCTT 976  
 Qy 961 GGGCTGGCTCAAGAGAGACTCTGCGGTGCTGTGCTGAGACTGAGAGAAAGAGTGA 1020  
 Db 977 GGGCTGGCTCAAGAGAGACTCTGCGGTGCTGTGCTGAGACTGAGAGAAAGAGTGA 1036  
 Qy 1021 GTTGTGCTTCCAGTATTAATCACTTACAGAGTAAATATTTTCAACTACAGTCTGGTG 1080  
 Db 1037 GTTGTGCTTCCAGTATTAATCACTTACAGAGTAAATATTTTCAACTACAGTCTGGTG 1096  
 Qy 1081 ATGAGCTGACGCGAGAACCTTCTTACCACTGGGCTGCAATTCATCGAGAGTAAAG 1140  
 Db 1097 ATGAGCTGACGCGAGAACCTTCTTACCACTGGGCTGCAATTCATCGAGAGTAAAG 1156  
 Qy 1141 GCCAGGCGCATCATGAGAGAGTATAGAGCTGTGAGGCCCTCAATATCATCAGGTC 1200  
 Db 1157 GCCAGGCGCATCATGAGAGAGTATAGAGCTGTGAGGCCCTCAATATCATCAGGTC 1216  
 Qy 1201 CTGAGCCATGAGAGAGGACACATCACTTTTGATCCAGCTGAGTGCCTGAGGCC 1260  
 Db 1217 CTGAGCCATGAGAGAGGACACATCACTTTTGATCCAGCTGAGTGCCTGAGGCC 1276  
 Qy 1261 AGTCTACTGATGATCATTAATCAAGTATTTCTTCCATCATCTCCACGAGACACATG 1320  
 Db 1277 AGTCTACTGATGATCATTAATCAAGTATTTCTTCCATCATCTCCACGAGACACATG 1336  
 Qy 1321 ACTGTGATGATCCAAAGAGATGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1380  
 Db 1337 ACTGTGATGATCCAAAGAGATGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1396  
 Qy 1381 GTTGTGCAACATGAGAAAGATGCTGCTTACGTC 1416  
 Db 1397 GTTGTGCAACATGAGAAAGATGCTGCTTACGTC 1432.

RESULT 3  
 AX006440 1860 bp DNA linear PAT 06-SEP-2000  
 LOCUS  
 DEFINITION Sequence 5 from Patent WO0004157.  
 AX006440  
 ACCESSION AX006440  
 VERSION AX006440.1 GI:9994575  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 Neeff, J.M., Peeters, D.C. and Pangalos, M.  
 Cloning and characterisation of novel mammalian peptidases  
 Patent: WO 0004157-A 5 27-JAN-2000;  
 JANSSEN PHARMACEUTICA NV (BE); NEEFS JEAN MARC EDMOND FERNAND (BE);  
 PETERS DANIELLE CELINE GEORGE (BE); PANGALOS MERLIS (GB)

FEATURES  
 source 1..1860  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 518 a 420 c 451 g 471 t  
 ORIGIN

Query Match 99.8%; Score 1412.8; DB 6; Length 1860;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 ATGAAATTCCTTATTCGATTTTGGTGGTGTGACCTTTATCCCTGCTGGG 60
Db 125 ATGAAATTCCTTATTCGATTTTGGTGGTGTGACCTTTATCCCTGCTGGG 184
Qy 61 AAAGCTATATGCAAGAAATGGCATCTTAAGAGACTTTTGAAGAAATTAAGAATA 120

```

Db 185 AAAGTATATGCAAGATGGATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATA 244  
 Qy 121 GCGAGCTGTGAGATGTTGCTAAAGCATCATAGCTGTTATGTTAAAGCCAG 180  
 Db 245 GCGAGCTGTGAGATGTTGCTAAAGCATCATAGCTGTTATGTTAAAGCCAG 304  
 Qy 181 AACAGATCTTATGAGCGATTTGGCACTTCTGTTGATCTGTGGAACCGAGCTGAGTGGC 240  
 Db 305 AACAGATCTTATGAGCGATTTGGCACTTCTGTTGATCTGTGGAACCGAGCTGAGTGGC 364  
 Qy 241 TCAGAAAGCTTAAAGAAAGCCATCCAAATATATGACAAAACCTGACAGCAAGATGGCTG 300  
 Db 365 TCAGAAAGCTTAAAGAAAGCCATCCAAATATATGACAAAACCTGACAGCAAGATGGCTG 424  
 Qy 301 GAGAAAGTTCACCTGAGAGCCAGTGAATATCCCACTGGAGAGAGGGAGAAAGATCACT 360  
 Db 425 GAGAAAGTTCACCTGAGAGCCAGTGAATATCCCACTGGAGAGAGGGAGAAAGATCACT 484  
 Qy 361 GTGATGCTGAGAGCCAGAAATTCATAGATAGCCATCTGGGTCTTGGAGAGCATTTGGG 420  
 Db 485 GTGATGCTGAGAGCCAGAAATTCATAGATAGCCATCTGGGTCTTGGAGAGCATTTGGG 544  
 Qy 421 ACTCTCCAGAAAGCATTAACAGCAGAGTTCTGTGTGACCTCTTTGATGATCTGAG 480  
 Db 545 ACTCTCCAGAAAGCATTAACAGCAGAGTTCTGTGTGACCTCTTTGATGATCTGAG 604  
 Qy 481 AGAAGGGCTCAGAAAGCAAGAGGAGATGTTGTTTAAACCACTTATCACTCACTAC 540  
 Db 605 AGAAGGGCTCAGAAAGCAAGAGGAGATGTTGTTTAAACCACTTATCACTCACTAC 664  
 Qy 541 TCAAGAGCGTGCATATACCAAGCAGAGGGCGGTGAGAGCTGCCAAGTGGGGCTTTG 600  
 Db 665 TCAAGAGCGTGCATATACCAAGCAGAGGGCGGTGAGAGCTGCCAAGTGGGGCTTTG 724  
 Qy 601 GCATCTCTCATTCGATCCGTGGCTCTCTTCTCATCTCAAGTCTCTCAACAGATTTAG 660  
 Db 725 GCATCTCTCATTCGATCCGTGGCTCTCTTCTCATCTCAAGTCTCTCAACAGATTTAG 784  
 Qy 661 GAATACCGAGATGGGCTGCCAAATTCACACAGCTGATTAAGTGAAGATGCAAA 720  
 Db 785 GAATACCGAGATGGGCTGCCAAAGTTCACACAGCTGATTAAGTGAAGATGCAAA 844  
 Qy 721 ATGATGTCAGAAATGGCTTCTCATAGGATTCAAAATTTGATTCAGCTTAAAGTGGGCGCA 780  
 Db 845 ATGATGTCAGAAATGGCTTCTCATAGGATTCAAAATTTGATTCAGCTTAAAGTGGGCGCA 904  
 Qy 781 AAGACTTACCAAGATAGTGAATCTTCTCAACACTGTAGCAGAGATCACTGGAGCAAAATAT 840  
 Db 905 AAGACTTACCAAGATAGTGAATCTTCTCAACACTGTAGCAGAGATCACTGGAGCAAAATAT 964  
 Qy 841 CCAGAAAGCTTGTATCTGTAGTGAATCTGACAGCTGGAGTGGGCGAGGCTGCC 900  
 Db 965 CCAGAAAGCTTGTATCTGTAGTGAATCTGACAGCTGGAGTGGGCGAGGCTGCC 1024  
 Qy 901 ATGATGATGCGCGGTGAGAGCTTTATATCAAGGAAGCACTCTCACTTATTAAGATCTT 960  
 Db 1025 ATGATGATGCGCGGTGAGAGCTTTATATCAAGGAAGCACTCTCACTTATTAAGATCTT 1084  
 Qy 961 GGGCTGCTCCAAAGAGAGACTTCTGCGGTGCTCTGAGCTGCAGAAAGCAAGTGGCA 1020  
 Db 1085 GGGCTGCTCCAAAGAGAGACTTCTGCGGTGCTCTGAGCTGCAGAAAGCAAGTGGCA 1144  
 Qy 1021 GTTGTGCTCCGATATTAATCAAGTATCAAGAAATTAATTTCCAACTACAGTCTGTG 1080  
 Db 1145 GTTGTGCTCCGATATTAATCAAGTATCAAGAAATTAATTTCCAACTACAGTCTGTG 1204  
 Qy 1081 ATGAGATCTGAGAGCAAGCTTCTTACCACTGGGCTGCAATTCAGTGGAGTAAAG 1140  
 Db 1205 ATGAGATCTGAGAGCAAGCTTCTTACCACTGGGCTGCAATTCAGTGGAGTAAAG 1264  
 Qy 1141 GCCAGAGGCATCATGAGAGAGTTATGAGCCTGTGAGAGCCCTCAATATCACTCAGGTC 1200  
 Db 1265 GCCAGAGGCATCATGAGAGAGTTATGAGCCTGTGAGAGCCCTCAATATCACTCAGGTC 1324

Qy 1201 CTGAGCCATGAGAGAGGAGAGCATCACTTTGATCCAAAGTGGAGTCCCTGAGACC 1260  
 Db 1325 CTGAGCCATGAGAGAGGAGAGCATCACTTTGATCCAAAGTGGAGTCCCTGAGACC 1384  
 Qy 1261 AGTCTACTGATGATCTTATACCAAGTATTTCTTTCCATGATCACTCCAGAGACCATG 1320  
 Db 1385 AGTCTACTGATGATCTTATACCAAGTATTTCTTTCCATGATCACTCCAGAGACCATG 1444  
 Qy 1321 ACTGTCTGATGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTAT 1380  
 Db 1445 ACTGTCTGATGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTAT 1504  
 Qy 1381 GTTGTGAGAGCATGAGAAAGATCTGCTAGTCC 1416  
 Db 1505 GTTGTGAGAGCATGAGAAAGATCTGCTAGTCC 1540  
 RESULT 4  
 BC020689  
 LOCUS  
 DEFINITION Homo sapiens, plasma glutamate carboxypeptidase, clone MGC:22418  
 IMAGE:4251802, mRNA, complete cds.  
 ACCESSION BC020689  
 VERSION BC020689.1 GI:18088383  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 1928)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 COMMENT Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC/DCTD/DTF  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>  
 Series: IRAL Plate: 35 Row: F Column: 13  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7706386.  
 FEATURES  
 location/Qualifiers  
 1..1928  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="LocustID:10404"  
 /db\_xref="taxon:9606"  
 /clone="MGC:22418 IMAGE:4251802"  
 /tissue\_type="Skin, melanocytic melanoma, high MDR."  
 /clone\_id="NIH\_MGC\_62"  
 /lab\_host="DH10B"  
 /note="Vector: pDNR-LIB"  
 167..1585  
 /codon\_start=1  
 /product="plasma glutamate carboxypeptidase"  
 /protein\_id="AAH20689.1"  
 /db\_xref="GI:18088384"  
 /translation="MKFLIPAFPGVHLILSGSKAIKNGISKRTFEIKKEIASCG DVAKAINTLAVGAQNSYERLALIVDTVGPRISGSKNEKAIQIWMYQNLQDGLK

```

VHLEPVRI PHERGESBAMLEPRHIALGLSSIGTPEGTAEVLVTSFDELQ
RASEBARKIVVNOPIYINSTRVOTQAVENAKGALASILRSVASFISPHG
IOBXOGBPKI PIACITVEDAKMSRMAISIKIYIOKMGAKTYPTDSFNVAET
ISXKPEQVVLVSHLDMDYGDMDGGGAFISWELSLIKDGLAEKRTLRVMT
ABROGGAFOYVOLHAKVNI SNYSIWMESAGFLPGLOPSTSEKRAIMEVMSLT
QPINIYVLSHGBSTDIINFMIQAGVPASLIDLYKFFPHSHGDMTMDPKONNV
AAAVWAVSVTVVMDMEMLPRS"
BASE COUNT      554 a      435 c      466 g      473 t
ORIGIN
Query Match      99.8%; Score 1412.8; DB 9; Length 1928;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 ATGAAATCTTATCTTGCGATTTTGGTGTGTTACCTTTATCCGTGTGGG 60
DB      167 ATGAAATCTTATCTTGCGATTTTGGTGTGTTACCTTTATCCGTGTGGG 226
QY      61 AAAGCTATGCAAGAAATGGCATCTTAAGAGACTTTTGAAGAAATMAAGAAATA 120
DB      227 AAAGCTATGCAAGAAATGGCATCTTAAGAGACTTTTGAAGAAATMAAGAAATA 286
QY      121 GCCAGCTGTGAGATGTTGCTTAAAGCAATCACTAGCTTTATGTTAAAGCCAG 180
DB      287 GCCAGCTGTGAGATGTTGCTTAAAGCAATCACTAGCTTTATGTTAAAGCCAG 346
QY      181 AACGATCTATGAGGATGGCACTTCTGTGTGATCTGTGGAGCCAGCTAGTGGC 240
DB      347 AACGATCTATGAGGATGGCACTTCTGTGTGATCTGTGGAGCCAGCTAGTGGC 406
QY      241 TCCAGAACTAGAAAAAGCCATCCAAATTATGTACCAAAAAGCTGACAGAAATGGGCTG 300
DB      407 TCCAGAACTAGAAAAAGCCATCCAAATTATGTACCAAAAAGCTGACAGAAATGGGCTG 466
QY      301 GAGAAATTTCACTGAGGCCAGTGAATACCCCACTGGAGAGAGGAGAAATCACT 360
DB      467 GAGAAATTTCACTGAGGCCAGTGAATACCCCACTGGAGAGAGGAGAAATCACT 526
QY      361 GTGTGCTGGAGCCAAATTCATTAAGTACCTCCGGGTCTTGGAGAGCAATGGG 420
DB      527 GTGTGCTGGAGCCAAATTCATTAAGTACCTCCGGGTCTTGGAGAGCAATGGG 586
QY      421 ACTCTCTCAGAAAGCAATTAAGCAAGATTTCTGTGTGATCTTTCATGAACTGAG 480
DB      587 ACTCTCTCAGAAAGCAATTAAGCAAGATTTCTGTGTGATCTTTCATGAACTGAG 646
QY      481 AGAAGGCTCTCAGAAAGCAAGATTTGTTTAAACCACTTACATCACTAC 540
DB      647 AGAAGGCTCTCAGAAAGCAAGATTTGTTTAAACCACTTACATCACTAC 706
QY      541 TCAAGAGCGGTGCAATTCAGAACGAGGGCGGTGGAAGCTGCCAAGTGGGGCTTGG 600
DB      707 TCAAGAGCGGTGCAATTCAGAACGAGGGCGGTGGAAGCTGCCAAGTGGGGCTTGG 766
QY      601 GCATCTCTCATTCGATCCGTGGCTCTCTTCTCATCTCAAGTCTCTCAACAGATTCAG 660
DB      767 GCATCTCTCATTCGATCCGTGGCTCTCTTCTCATCTCAAGTCTCTCAACAGATTCAG 826
QY      661 GAATACAGAGATGGCGCCCAAAATTCCAACAGCTGTATTAAGTGAAGATGCAAA 720
DB      827 GAATACAGAGATGGCGCCCAAAATTCCAACAGCTGTATTAAGTGAAGATGCAAA 886
QY      721 ATGATGTCAGAAATGGCTTCTCATGAGATCAAAATTTGATTCAGCTTAAGATGGGGCA 780
DB      887 ATGATGTCAGAAATGGCTTCTCATGAGATCAAAATTTGATTCAGCTTAAGATGGGGCA 946
QY      781 AAGACCTAACCGAGATCTGATTTCTTCAACACTGTAGACAGATCACTGGAGCAAAATAT 840
DB      947 AAGACCTAACCGAGATCTGATTTCTTCAACACTGTAGACAGATCACTGGAGCAAAATAT 1006
QY      841 CCAGAACAGGTTGTACTGTGATGAGCATCTGGAAGCTGGAGTGTGGCAGGGTGGC 900
DB      1007 CCAGAACAGGTTGTACTGTGATGAGCATCTGGAAGCTGGAGTGTGGCAGGGTGGC 1066

```

```

QY      901 ATGATGATGGCGGTGAGCGCTTTATATCATGGGAAGCACTCTCACTTATTAAGATCTT 960
DB      1067 ATGATGATGGCGGTGAGCGCTTTATATCATGGGAAGCACTCTCACTTATTAAGATCTT 1126
QY      961 GGGCTGCGTCCAAAGAGACTCTGCGGCTGTCTGTGACTGACAGAAACAGGTGA 1020
DB      1127 GGGCTGCGTCCAAAGAGACTCTGCGGCTGTCTGTGACTGACAGAAACAGGTGA 1186
QY      1021 GTTGGGCTTCAGATTAATCATGATACAAAGGTAATTTTCCAACTACGTGGTG 1080
DB      1187 GTTGGGCTTCAGATTAATCATGATACAAAGGTAATTTTCCAACTACGTGGTG 1246
QY      1081 ATGAGCTGACCGCAGGAACCTTTTACCACTGAGGCTGCAATTCATCTGACAGTAAAG 1140
DB      1247 ATGAGCTGACCGCAGGAACCTTTTACCACTGAGGCTGCAATTCATCTGACAGTAAAG 1306
QY      1141 GCCAGGGCATCATGAGAGAGTTATGAGGCTGTGAGGCCCTCAATATCACTAGTTC 1200
DB      1307 GCCAGGGCATCATGAGAGAGTTATGAGGCTGTGAGGCCCTCAATATCACTAGTTC 1366
QY      1201 CTGAGCCATGGAAGAGGACAGACATCACTTTTGAATCCAACTGAGAGTGGCTGAGCC 1260
DB      1367 CTGAGCCATGGAAGAGGACAGACATCACTTTTGAATCCAACTGAGAGTGGCTGAGCC 1426
QY      1261 AGTCTACTGATGACTTATACAGTATTTCTTCCATCACTCCACGAGACACCATG 1320
DB      1427 AGTCTACTGATGACTTATACAGTATTTCTTCCATCACTCCACGAGACACCATG 1486
QY      1321 ACTGTCAATGATCCAAAGCAGATGAATGTTGCTGCTGTGTTGGCTGTGTTCTTAT 1380
DB      1487 ACTGTCAATGATCCAAAGCAGATGAATGTTGCTGCTGTGTTGGCTGTGTTCTTAT 1546
QY      1381 GTTGTTCAGACATGGAAGAAATGCTGCTAGTTC 1416
DB      1547 GTTGTTCAGACATGGAAGAAATGCTGCTAGTTC 1582

RESULT 5
AF107834      1472 bp      mRNA      linear      PRI 12-JUL-1999
LOCUS      AF107834
DEFINITION      Homo sapiens clone LCH1 aminopeptidase mRNA, complete cds.
ACCESSION      AF107834
VERSION      AF107834.1 GI:5442029
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1. (bases 1 to 1472)
AUTHORS      Liu, C.H., Lin, B.Y. and Chang, L.Y.
TITLE      Cloning of the human aminopeptidase gene
JOURNAL      Unpublished
REFERENCE      2. (bases 1 to 1472)
AUTHORS      Liu, C.H., Lin, B.Y. and Chang, L.Y.
TITLE      Direct Submission
JOURNAL      Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia
SINICA, RM 433, 128, Yen-Chun-Yuan Road SEC 2, Taipei 11529, Taiwan
FEATURES
source
1. 1472
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="LCH1"
/lisue_type="liver"
/dev_stage="fetus"
54..1472
/note="secretory protein"
/codon_start=1
/product="aminopeptidase"
/protein_id="A043214.1"
/db_xref="GI:5442030"
/translaction="MKFLIPAFGVHLISLCSGKAIKNGISKRTFEEIKETIASCG

```

DVAKATINLAVYGAONRSYERLALVDYVPRLSGSKNLEKAIQIMYONIQODLEK  
 VHLIEPVRIPHEEREEASVMEPRIRHKAIIIGLSSIGTPEGITATLAVTSDELO  
 RASEBARKIVYNNOPRYNERTYOTQOAGVEAKGALSILRSVASETSYBPHG  
 IOETODVPKIPTACTIVEDKEMMSHSHGKITVIOIKMGAKTYPTDSENTYAEIT  
 GSKYREOVVLSGILDSMDVGOAGMDGGAFIEMWLSLKDGLPKRRLRLVLT  
 AEEGCVGAFOYQOLHKVNI SNYSIWMESDAGETPLGLOPTGSEKARAIEMEWLSL  
 OPLNITGVLSHGESEIDINFMIQAGVPGASLLDLIDYKFFPHSHQDWTWMDPKOMV  
 AAAYVAVSYVVAEMEELPRS"

BASE COUNT 411 a 325 c 381 g 355 t  
 ORIGIN  
 Query Match 99.5%; Score 1409.6; DB 9; Length 1472;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 1412; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAATTCCTTATCTTGGCAATTTTGGTGTGTCACCTTTATCCCTGTGCTGGG 60  
 DB 54 ATGAAATTCCTTATCTTGGCAATTTTGGTGTGTCACCTTTATCCCTGTGCTGGG 113  
 QY 61 AAAGCTATGCAAGATGGGATCTCTAAGAGGACTTTTGAAGAAATAAAGAAATA 120  
 DB 114 AAAGCTATGCAAGATGGGATCTCTAAGAGGACTTTTGAAGAAATAAAGAAATA 173  
 QY 121 GCCAGCTGTGAGATGTTGCTTAAAGCAATCACTAGCTGTTTATGTTAAAGCCAG 180  
 DB 174 GCCAGCTGTGAGATGTTGCTTAAAGCAATCACTAGCTGTTTATGTTAAAGCCAG 233  
 QY 181 AACGATCTTATGAGCCATTGGCACTTGTGTTGATCTTTGGAACCCAGCTGAGTGC 240  
 DB 234 AACGATCTTATGAGCCATTGGCACTTGTGTTGATCTTTGGAACCCAGCTGAGTGC 293  
 QY 241 TCCAGAACTTAAAGAAAGCCATCCAAATTAATGTAACCAAACTGCGACGAAATGGGCTG 300  
 DB 294 TCCAGAACTTAAAGAAAGCCATCCAAATTAATGTAACCAAACTGCGACGAAATGGGCTG 353  
 QY 301 GAGAAAGTTCACTGAGCCAGTGAAGTAACCCACTGCGAGAGGGGAGAAATCAAGCT 360  
 DB 354 GAGAAAGTTCACTGAGCCAGTGAAGTAACCCACTGCGAGAGGGGAGAAATCAAGCT 413  
 QY 361 GTGATGCTGAGCCAGAAATTCATTAAGTATACCATCTCGGCTTTGGAGACGATGGG 420  
 DB 414 GTGATGCTGAGCCAGAAATTCATTAAGTATACCATCTCGGCTTTGGAGACGATGGG 473  
 QY 421 ACTCCTCAGAAAGGATTAAGCAAGAAATTCGAGGAGTGAAGCTCTTTGATGAATCTGAG 480  
 DB 474 ACTCCTCAGAAAGGATTAAGCAAGAAATTCGAGGAGTGAAGCTCTTTGATGAATCTGAG 533  
 QY 481 AGAAGGGCTTCAGAAAGCAAGGAGAAATGTTGTTTAAACAACCTTATCATCACTAC 540  
 DB 534 AGAAGGGCTTCAGAAAGCAAGGAGAAATGTTGTTTAAACAACCTTATCATCACTAC 593  
 QY 541 TCAAGACGGTGCATATACCAAGCGAGGGGCGGTGAGAGCTGCCAAGTGGGGCTTTG 600  
 DB 594 TCAAGACGGTGCATATACCAAGCGAGGGGCGGTGAGAGCTGCCAAGTGGGGCTTTG 653  
 QY 601 GCATCTCTCATTCGATCCGTGGCTCTCTTCATCTAAGTCCCTCAACAGGATTCAG 660  
 DB 654 GCATCTCTCATTCGATCCGTGGCTCTCTTCATCTAAGTCCCTCAACAGGATTCAG 713  
 QY 661 GAATACAGAGATGGCGTGCACCAAAATTCACAAGCTGTATTAACGTTGAAGATGCAAA 720  
 DB 714 GAATACAGAGATGGCGTGCACCAAGTTCACAAGCTGTATTAACGTTGAAGATGCAAA 773  
 QY 721 ATGATGTCAGAAATGGCTTCTCATGAGGATCAAAATGTCATTCAGCTTAAAGATGGGGCA 780  
 DB 774 ATGATGTCAGAAATGGCTTCTCATGAGGATCAAAATGTCATTCAGCTTAAAGATGGGGCA 833  
 QY 781 AAGACCTAACCAAGATGATCTTCTTCAACCTGTAGACAGATCACTGGAGCAAAATAT 840  
 DB 834 AAGACCTAACCAAGATGATCTTCTTCAACCTGTAGACAGATCACTGGAGCAAAATAT 893  
 QY 841 CCAAGACAGGTTGTACTGTGATGAGCATCTGCAAGCTGGAGTGTGGCAGGGTGC 900

DB 894 CCAAGACAGGTTGTACTGTCTAGTGACATCTGACAGACTGGAGATGTTGGGCAAGGCTGC 953  
 QY 901 ATGATATATGGCGGTGAGGCTTTATATCAATGGGAACACTCTCACTATTAAAGATCTT 960  
 DB 954 ATGATATATGGCGGTGAGGCTTTATATCAATGGGAACACTCTCACTATTAAAGATCTT 1013  
 QY 961 GGGCTGCGTCCAAAGAGACTCTGCGCTGGTGTCTGAGACTGAGAAACAAGATGGA 1020  
 DB 1014 GGGCTGCGTCCAAAGAGACTCTGCGCTGGTGTCTGAGACTGAGAAACAAGATGGA 1073  
 QY 1021 GTTGTGCTCTTCCAGATTTATACGTTACACAGGTAATTTTCCACTACAGTCTGGT 1080  
 DB 1074 GTTGTGCTCTTCCAGATTTATACGTTACACAGGTAATTTTCCACTACAGTCTGGT 1133  
 QY 1081 ATGAGCTTACAGGAGAACTTCTTACCACTGGGCTGGAATTCATGAGCAAGTAAAG 1140  
 DB 1134 ATGAGCTTACAGGAGAACTTCTTACCACTGGGCTGGAATTCATGAGCAAGTAAAG 1193  
 QY 1141 GCCAGGGCCATCATGAGAGAGGTTATAGGCTGTGCAAGCCCTCAATATCACTCAGGTC 1200  
 DB 1194 GCCAGGGCCATCATGAGAGAGGTTATAGGCTGTGCAAGCCCTCAATATCACTCAGGTC 1253  
 QY 1201 CTGAGCCATGAGAGAGGAGACAGACATCACTTTGGATCCAGCTGAGTGCCTGAGCC 1260  
 DB 1254 CTGAGCCATGAGAGAGGAGACAGACATCACTTTGGATCCAGCTGAGTGCCTGAGCC 1313  
 QY 1261 AGTCTACTTATGATCACTTATACAGTATTTCTTCCATCATCTCCACGAGACACATG 1320  
 DB 1314 AGTCTACTTATGATCACTTATACAGTATTTCTTCCATCATCTCCACGAGACACATG 1373  
 QY 1321 ACTGTCAATGATCCAAAGACAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1380  
 DB 1374 ACTGTCAATGATCCAAAGACAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1433  
 QY 1381 GTTGTCCAGACATGAGAAATGCTGCTTACGTTCC 1416  
 DB 1434 GTTGTCCAGACATGAGAAATGCTGCTTACGTTCC 1469

RESULT 6  
 LOCUS BD127520 1860 bp DNA linear PAT 18-SEP-2002  
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
 ACCESSION BD127520.1 GI:23222465  
 VERSION BD127520.1 GI:23222465  
 KEYWORDS UP 2002017375-A/2951.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1860)  
 Ota,T., Nishikawa,T., Isogai,T., Hayaishi,K., Ishii,S., Kawai,Y.,  
 Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and  
 Koga,H.  
 TITLE Primer for synthesizing full-length cDNA and use thereof  
 JOURNAL Patent: JP 2002017375-A 2951 22-JAN-2002;  
 HELIX RESEARCH INSTITUTE  
 OS Homo sapiens (human)  
 PN UP 2002017375-A/2951  
 PD 22-JAN-2002  
 PF 07-JUL-2000 JP 2000253172  
 PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO  
 PI ISHII,  
 PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI  
 SHINICHI KOJIMA,  
 PI TETSUJI OTSUKI, HISASHI KOGA  
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/PC  
 10', C12P21/02, C1201/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
 Primer for synthesizing full-length cDNA and use thereof FH Key  
 Location/Qualifiers  
 FT CDS (129)..(1544).

FEATURES  
Source Location/Qualifiers  
1..1860  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 515 a 423 c 451 g 471 t  
ORIGIN

Query Match 99.4%; Score 1408; DB 6; Length 1860;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1411; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 ATGAATTCCTTATCTTGGATTTTGGGAGTTCACCTTTATCCCTGCTCTGG 60  
129 ATGAATTCCTTATCTTGGATTTTGGGAGTTCACCTTTATCCCTGCTCTGG 188  
61 AAACCTATATGCAAGAAATGGCACTCTTAAGAGACTTTTGAAGAAATGAAGAAATA 120  
189 AAACCTATATGCAAGAAATGGCACTCTTAAGAGACTTTTGAAGAAATGAAGAAATA 248  
121 GCCAGCTGTGAGATGTTGTCTTAAGCAATCACTAGCTGTTTATGTTAAAGCCGAG 180  
249 GCCAGCTGTGAGATGTTGTCTTAAGCAATCACTAGCTGTTTATGTTAAAGCCGAG 308  
181 AACAGATCTTATGAGCGATTGGCACTTGTGTTGATCTGTTGAGCCGAGCTGAGTGC 240  
309 AACAGATCTTATGAGCGATTGGCACTTGTGTTGATCTGTTGAGCCGAGCTGAGTGC 368  
241 TCCAGAACTTAAAGAAAGCCATCCAAATTTATGCCAAACCTGCGAGCAAGATGGGCTG 300  
369 TCCAGAACTTAAAGAAAGCCATCCAAATTTATGCCAAACCTGCGAGCAAGATGGGCTG 428  
301 GAGAAAGTTTCACTGAGACCGAGTGAATACCCCACTGGGAGAGGGGAGAAATCAGCT 360  
429 GAGAAAGTTTCACTGAGACCGAGTGAATACCCCACTGGGAGAGGGGAGAAATCAGCT 488  
361 GTGATGCTGAGCCAGAAATTCATTAAGTATAGCCATCTCGGTCCTTGGCAGCAGATTGGG 420  
489 GTGATGCTGAGCCAGAAATTCATTAAGTATAGCCATCTCGGTCCTTGGCAGCAGATTGGG 548  
421 ACTCTCTCAAGAGGATTTACGACAGAGTTTGTGTTGATCTCTTTGATGAATCTGAG 480  
549 ACTCTCTCAAGAGGATTTACGACAGAGTTTGTGTTGATCTCTTTGATGAATCTGAG 608  
481 AGAAGGGGCTCAGAGAGCAAGAGGAGATGTTGTTTAAACCAACCTTATCACTAC 540  
609 AGAAGGGGCTCAGAGAGCAAGAGGAGATGTTGTTTAAACCAACCTTATCACTAC 668  
541 TCAAGAGCGGTGCAATACCGAAGCGAGGGCGGTGGAAGCTGCCAAGGTGGGGCTTTG 600  
669 TCAAGAGCGGTGCAATACCGAAGCGAGGGCGGTGGAAGCTGCCAAGGTGGGGCTTTG 728  
601 GCATCTCTCATTCGATCCGTCGCTCTTCTTCATCTACAGTCTCTCACAGGATTTAG 660  
729 GCATCTCTCATTCGATCCGTCGCTCTTCTTCATCTACAGTCTCTCACAGGATTTAG 788  
661 GAATACAGAGTGGGTCGCCAAATTCACACAGCTGTATTAAGGTGAAGATGCGAGAA 720  
789 GAATACAGAGTGGGTCGCCAAATTCACACAGCTGTATTAAGGTGAAGATGCGAGAA 848  
721 ATGATGTCAAGAAATGCTTCTCATGAGATCAAAATTTGATTCAGCTAAAGTGGGGCA 780  
849 ATGATGTCAAGAAATGCTTCTCATGAGATCAAAATTTGATTCAGCTAAAGTGGGGCA 908  
781 AAGACCTTACCAAGATGATGATTTCTTCAACCTGAGCAGAGATCACTGGAGCAAAATAT 840  
909 AAGACCTTACCAAGATGATGATTTCTTCAACCTGAGCAGAGATCACTGGAGCAAAATAT 968  
841 CCAGAAACAGGTTGATCTGTCAGTGAAGATCTGCAACACTGGAGATGTTGGCAGAGTGC 900  
969 CCAGAAACAGGTTGATCTGTCAGTGAAGATCTGCAACACTGGAGATGTTGGCAGAGTGC 1028  
901 ATGATGTATGCGGTGAGGCTTTATATCATGGAAGCACTCACTTATTAAGATCTT 960

Db 1029 ATGATGTATGCGGTGAGGCTTTATATCATGGAGACACTCACTTATTAAGATCTT 1088  
Qy 961 GCGCTGCGTCCAAAGAGACCTCTGCGGCTGCTCTGAGATCGAAGAAAGAGTGA 1020  
Db 1089 GGGCTGCGTCCAAAGAGACCTCTGCGGCTGCTCTGAGATCGAAGAAAGAGTGA 1148  
Qy 1021 GTTGGTGCCTCCAGATTTATCATGTTACACAAAGGTAATTTCCAACTACAGTGGTG 1080  
Db 1149 GTTGGTGCCTCCAGATTTATCATGTTACACAAAGGTAATTTCCAACTACAGTGGTG 1208  
Qy 1081 ATGAGTCTGACGAGAACTTCTTACCACTGGCTGCAATTCATCTGAGAGAAAG 1140  
Db 1209 ATGAGTCTGACGAGAACTTCTTACCACTGGCTGCAATTCATCTGAGAGAAAG 1268  
Qy 1141 GCCAGGCGCATGAGAGAGGTTATGAGCTGCTGCAAGCCCTGATATCATCTAGTGC 1200  
Db 1269 GCCAGGCGCATGAGAGAGGTTATGAGCTGCTGCAAGCCCTGATATCATCTAGTGC 1328  
Qy 1201 CTGAGCCATGAGAAAGGACAGATCACTTTTGAATCCAAAGTGGAGTGCCTGGAGCC 1260  
Db 1329 CTGAGCCATGAGAAAGGACAGATCACTTTTGAATCCAAAGTGGAGTGCCTGGAGCC 1388  
Qy 1261 AGTCTACTGATGACTTATACAGATTTCTTCTTCCATCACTCCACGAGACCATG 1320  
Db 1389 AGTCTACTGATGACTTATACAGATTTCTTCTTCCATCACTCCACGAGACCATG 1448  
Qy 1321 ACTGTCAATGATCCAAAGCAGATGAATGTTGCTGCTGCTGTTGGGCTGTTCTTAT 1380  
Db 1449 ACTGTCAATGATCCAAAGCAGATGAATGTTGCTGCTGCTGTTGGGCTGTTCTTAT 1508  
Qy 1381 GTTGTGACAGCATGAGAAAGTCTGCTAGTGC 1416  
Db 1509 GTTGTGACAGCATGAGAAAGTCTGCTAGTGC 1544

RESULT 7  
AK075132  
LOCUS  
DEFINITION  
AK075132 1860 bp mRNA linear PRI 03-SEP-2002  
Homo sapiens cDNA FLJ90651 fis, clone PLACE100482, moderately similar to Rat1us norvegicus hematopoietic lineage switch 2 related protein (H1ez-rp) mRNA.  
AK075132  
ACCESSION  
AK075132.1 GI:22761022  
VERSION  
Oligo capping, fis (full insert sequence).  
KEYWORDS  
Homo sapiens (human)  
SOURCE  
Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Isegai,T., Ota,T., Nishikawa,T., Hayaishi,K., Otsuki,T., Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S., Kawai,Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Kojima,S., Nagahari,K., Masuno,Y., Ono,T., Okano,K., Yoshikawa,Y., Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and Ninomiya,K.  
NEO human cDNA sequencing project  
TITLE  
NEO human cDNA sequencing project  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 1860)  
Isegai,T. and Otsuki,T.  
AUTHORS  
Isegai,T. and Otsuki,T.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (25-MAR-2002) Takao Isegai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
COMMENT  
NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).  
FEATURES  
Source Location/Qualifiers  
1..1860



```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1004482"
/tissue_type="placenta"
/clone_id="PLACE1"
/notes="cloning vector: PME18SFL3"
129.1547
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC11423.1"
/db_xref="GI:22761023"
/translation="MKFLIFAFPGVHLLSLCSKAIKNGISKREIEIEBIAISG
DVAKIIMLAYGKONRSYERLALVDVTPRLSGKYLEKAIQIMYQNDQGLEK
VHLEPVRIPHEWGESAVMLERPHKIAIIGLSSIGTPEPGIAEVLVYSFDELQ
RASEARKIIVYNOPYINVSRTVOYRTOGAVEAKVGLASLIRSVASPSIYSBHT
IOERODGPKIPIACITTEDBEMSRMSHGKIKYIOLKMGAKTYPDTSNNTVAET
GSKYPEOVNLVSHLSDVDVGAMDVGGAFLSWEALSLIKDGLRKRRLRLVLTW
AEBOGVGAFQYQLHKVNIISYSLVMSDGTPLPTGLQFTGSBKAAIIVEVMSLL
OPLNITOVLSHGEFTDINFMIQAGVPGASLDDLYKYFFHSHSDTWTMDPKOMVY
AAAVAVVSYVVADEEMLPSS"

BASE COUNT      515 a      423 c      451 g      471 t
ORIGIN
Query Match      99.4%; Score 1408; DB 9; Length 1860;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1411; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 ATGAATTCCTTATCTTGGCATTTCGCGTGTGTCACCTTTATCCCTGCTCTGGG 60
Db 129 ATGAATTCCTTATCTTGGCATTTCGCGTGTGTCACCTTTATCCCTGCTCTGGG 188
Oy 61 AAAGCTATATGCAAGATGCGATCTCTAAGAGGACCTTTTGAAGAAATAAAGAAATA 120
Db 189 AAAGCTATATGCAAGATGCGATCTCTAAGAGGACCTTTTGAAGAAATAAAGAAATA 248
Oy 121 GCCAGCTGTGAGATGTGCTTAAGCAATTCATCACTAGCTGTTATGTAAAGCCAG 180
Db 249 GCCAGCTGTGAGATGTGCTTAAGCAATTCATCACTAGCTGTTATGTAAAGCCAG 308
Oy 181 AACAGATCTTATGACGATGGCACTTCTGTGTGATCTGTTGACCCAGACTGAGTGGC 240
Db 309 AACAGATCTTATGACGATGGCACTTCTGTGTGATCTGTTGACCCAGACTGAGTGGC 368
Oy 241 TCCAGAACCTAGAAAGCAATCCAAATTAATGTACCAAAACCTGACAGAAATGGGGT 300
Db 369 TCCAGAACCTAGAAAGCAATCCAAATTAATGTACCAAAACCTGACAGAAATGGGGT 428
Oy 301 GAGAAAGTTCACTGTGAGCCAGTGAATATCCCACTGGGAGAGGGAGAAATCAGCT 360
Db 429 GAGAAAGTTCACTGTGAGCCAGTGAATATCCCACTGGGAGAGGGAGAAATCAGCT 488
Oy 361 GTGATGCTGAGCCAGAAATTCATTAAGATAGCCATCTGGGCTTTGGCAGCAGATTGGG 420
Db 489 GTGATGCTGAGCCAGAAATTCATTAAGATAGCCATCTGGGCTTTGGCAGCAGATTGGG 548
Oy 421 ACTGCTCAGAGAGGATTAACAGCAAGTCTGCGTGTGAGACCTCTTTGATGAATACTGAG 480
Db 549 ACTGCTCAGAGAGGATTAACAGCAAGTCTGCGTGTGAGACCTCTTTGATGAATACTGAG 608
Oy 481 AGAAGGCGCTCAGAAAGAGAGAGATTTGTTTAAACCAACTTACATCACTAC 540
Db 609 AGAAGGCGCTCAGAAAGAGAGAGATTTGTTTAAACCAACTTACATCACTAC 668
Oy 541 TCAAGAGCGTGCAATACCGAAACGAGGGGGGGGTGAAGCTGCCAGGTGGGGCTTTG 600
Db 669 TCAAGAGCGTGCAATACCGAAACGAGGGGGGGGTGAAGCTGCCAGGTGGGGCTTTG 728
Oy 601 GCATCTTCATTCGATCCGTGGCTCTTCTTCATCTTACAGTCCCTCAACAGATTTTAC 660
Db 729 GCATCTTCATTCGATCCGTGGCTCTTCTTCATCTTACAGTCCCTCAACAGATTTTAC 788
Oy 661 GAATACAGAGATGGCGTCCCAAAATTCACAACAGCCTGTATTAAGTGAAGATGCAGAA 720

```

```

|||||
Db 789 GAATACAGAGATGGCGTCCCAAAATTCACAACAGCCTGTATTAAGTGAAGATGCAGAA 848
Oy 721 ATGATGTCAGAAATAGGCTTCTCATAGGATCAAAATTTGCTATTCAGTAAAGATGGGGCA 780
Db 849 ATGATGTCAGAAATAGGCTTCTCATAGGATCAAAATTTGCTATTCAGTAAAGATGGGGCA 908
Oy 781 AAGACTTACCAGATATCTGATTCCTTCAACACTGTACAGAGATCATCTGGAGCAATAT 840
Db 909 AAGACTTACCAGATATCTGATTCCTTCAACACTGTACAGAGATCATCTGGAGCAATAT 968
Oy 841 CCAGAACAGGTTTACTGTGTCAGTGAACATCTGACAGCTGGAGTGTGGCAGGGTGC 900
Db 969 CCAGAACAGGTTTACTGTGTCAGTGAACATCTGACAGCTGGAGTGTGGCAGGGTGC 1028
Oy 901 ATGATATAGCGCGGTGAGAGCTTTATATCATATGGAACACCTTCACCTAATTAAGATCT 960
Db 1029 ATGATATAGCGCGGTGAGAGCTTTATATCATATGGAACACCTTCACCTAATTAAGATCT 1088
Oy 961 GGGCTGCGTCCAAAGAGAGACTCTGCGGCTGCTGTGACCTGAGACAGAAACAAGTGA 1020
Db 1089 GGGCTGCGTCCAAAGAGAGACTCTGCGGCTGCTGTGACCTGAGACAGAAACAAGTGA 1148
Oy 1021 GTTGTGCTCTTCCAGATATTAACAGTTACACAGGTAAATATTTCCAACTACAGTCTGTG 1080
Db 1149 GTTGTGCTCTTCCAGATATTAACAGTTACACAGGTAAATATTTCCAACTACAGTCTGTG 1208
Oy 1081 ATGAGATCTGACGAGAGAACCTTTTACCACTGAGGCTGCAATTCATCTGGAGAGAAAG 1140
Db 1209 ATGAGATCTGACGAGAGAACCTTTTACCACTGAGGCTGCAATTCATCTGGAGAGAAAG 1268
Oy 1141 GCCAGGCGCATCATGAGAGAGGTTATGAGCTGTGACGCCCTCAATATCATCTCAGGTC 1200
Db 1269 GCCAGGCGCATCATGAGAGAGGTTATGAGCTGTGACGCCCTCAATATCATCTCAGGTC 1328
Oy 1201 CTGAGCCATGAGAGAGGAGACAGACATCACTTTTGAATCCAACTGAGTGCCTGAGCC 1260
Db 1329 CTGAGCCATGAGAGAGGAGACAGACATCACTTTTGAATCCAACTGAGTGCCTGAGCC 1388
Oy 1261 AGTTACTTATGATCATTAACAGATATTTCTTCCATCACTCTCCACGAGAACCAATG 1320
Db 1389 AGTTACTTATGATCATTAACAGATATTTCTTCCATCACTCTCCACGAGAACCAATG 1448
Oy 1321 ACTGTCAATGATCCAAAGAGAGATGATGTTGCTGCTGCTGTTGGGCTGTTCTTAT 1380
Db 1449 ACTGTCAATGATCCAAAGAGAGATGATGTTGCTGCTGCTGTTGGGCTGTTCTTAT 1508
Oy 1381 GTTGTGCAAGCATGAGAAAGAAATGCTGCTAGTCC 1416
Db 1509 GTTGTGCAAGCATGAGAAAGAAATGCTGCTAGTCC 1544

RESULT 8
AF119386 1794 bp mRNA linear PRI 20-MAY-1999
LOCUS Homo sapiens blood plasma glutamate carboxypeptidase precursor
DEFINITION (BCCP) mRNA, complete cds.
ACCESSION AF119386
VERSION AF119386.1 GI:4877697
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1794)
AUTHORS Gingras, R., Richard, C., Bl-Alfy, M., Morales, C.R., Potier, M. and
Pehzetsky, A.V.
Purification, cDNA cloning, and expression of a new human blood
plasma glutamate carboxypeptidase homologous to
N-acetyl-aspartyl-alpha-glutamate
carboxypeptidase/prostate-specific membrane antigen
JOURNAL J. Biol. Chem. 274 (17), 11742-11750 (1999)
MEDLINE 99223495

```



PUBMED 10206990  
REFERENCE 2 (bases 1 to 1794)  
AUTHORS Gingras, R., Richard, C., El-Alfy, M., Morales, C. R., Potier, M. and Pehezhetsky, A. V.  
TITLE Direct Submission  
JOURNAL Submitted (12-JUN-1999) Medical Genetics, Sainte-Justine Hospital, Montreal University, 3175 Cote Sainte-Catherine, Montreal, QU H3T 1C5, Canada

FEATURES  
source  
1. 1794  
/location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8q22.2; between DBS257 and DBS270"  
1. 1794  
/gene="PGCP"  
25. 1650  
/EC number="3.4.17.21"  
/gene="PGCP"  
/EC number="3.4.17.21"  
/function="endopeptidase"  
/note="member of M28 peptidase family; similar to glutamate carboxypeptidase II and to N-acetyl-alpha-glutamate carboxypeptidase (NALADase); similar to sequences with GenBank Accession Numbers R18560 and G07404; prostate-specific membrane antigen (PSMA)"  
/codon\_start=1  
/product="blood plasma glutamate carboxypeptidase precursor"  
/protein\_id="AAD31438.1"  
/db\_xref="GI:4877698"  
/translation="MKFLIFAFGFGVHLISGSKAIKNGISKRTPEIKIEIASG DVAKILINAVYGAONRSYERLALVDYTPRSGSKNLEKAIQIMYONLQDGLK VHEPVAIPMERGESAVMEPRRIKILIGLSSITGPEGITAEVLVTSFDEIQ RISEARAKIVINQPIYINSTRVQYRTQGAENAKVGLASLIRVASFISYPTG IOEQDQVPKIPACITVEDAEKMSRASHGKIIVLQKGAQKTPIDSPNTVAETG ASKYPEQVVLVSGHDSMDVQCGAMDGGGAFISMEALSILKIDGLRPKETRLVLT ABECKGVGARFOYOLHKVNISNYSLVMSDSAGTLPGLQDTSSEKARALMEEVMSLL OPLNITOVLSHGEGTDINFWIOAGVPGASLIDLCYKFFPHSHGDMYTWDERMWL LILFGLPLMLLOTWKCCUCGPRNSKKTETSCFPGLIGLQMLKTPHLITTSNSSK HNSISCFLLSLFSLKPSDSRKNHSPLEPTT"  
25. 96  
/gene="PGCP"  
157. 1647  
/gene="PGCP"  
/product="blood plasma glutamate carboxypeptidase"  
205. 207  
/gene="PGCP"  
/note="putative; glycosylation site"  
559. 561  
/gene="PGCP"  
/note="putative; glycosylation site"  
misc\_feature  
/note="putative; glycosylation site"  
1081. 1083  
/gene="PGCP"  
/note="putative; glycosylation site"  
misc\_feature  
1090. 1092  
/gene="PGCP"  
/note="putative; glycosylation site"  
misc\_feature  
1210. 1212  
/gene="PGCP"  
/note="putative; glycosylation site"  
misc\_feature  
/note="putative; glycosylation site"  
BASE COUNT 522 a 391 c 422 g 457 t 2 others  
ORIGIN

Query Match 98.8%; Score 139.2; DB 9; Length 1794;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1412; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 ATGAATTCCTTATCTTGGATTTTGGTGTGTTACCTTTATCCTGTGCTTGGG 60  
DB 25 ATGAATTCCTTATCTTGGATTTTGGTGTGTTACCTTTATCCTGTGCTTGGG 84

QY 61 AAAGCTATATGAGAGATGGCATCTCTAAGAGAGACTTTTGAAGAAATTAAGAGAAATA 120  
DB 85 AAAGCTATATGAGAGATGGCATCTCTAAGAGAGACTTTTGAAGAAATTAAGAGAAATA 144  
QY 121 GCCAGCTGTGAGATGTTGCTTAAAGCATATCAACTAGCTGTTTATGTTAAAGCCAG 180  
DB 145 GCCAGCTGTGAGATGTTGCTTAAAGCATATCAACTAGCTGTTTATGTTAAAGCCAG 204  
QY 181 AACGATCTATGAGGATGGGACTCTGTTGATATCTGTGGAGCCAGCTGAGTGGC 240  
DB 205 AACGATCTATGAGGATGGGACTCTGTTGATATCTGTGGAGCCAGCTGAGTGGC 264  
QY 241 TCAGAGAACCTAGAAAAAGCCATCCAAATTATGATCCAAAACTGACAGAAATGGCTG 300  
DB 265 TCAGAGAACCTAGAAAAAGCCATCCAAATTATGATCCAAAACTGACAGAAATGGCTG 324  
QY 301 GAGAAAGTTCACTGAGAGCAGTGAATATCCCACTGGGAGAGGGAGAAATCAAGCT 360  
DB 325 GAGAAAGTTCACTGAGAGCAGTGAATATCCCACTGGGAGAGGGAGAAATCAAGCT 384  
QY 361 GTGATGCTGAGGCAAGAAATTCATTAAGATAGCCATCCTGGGCTTGGCAGAGATTGGG 420  
DB 385 GTGATGCTGAGGCAAGAAATTCATTAAGATAGCCATCCTGGGCTTGGCAGAGATTGGG 444  
QY 421 ACTCTCCAGAGGCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
DB 445 ACTCTCCAGAGGCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 504  
QY 481 AGAAGGCTTCAG 540  
DB 505 AGAAGGCTTCAG 564  
QY 541 TCAAG 600  
DB 565 TCAAG 624  
QY 601 GCATCTCTATGATGATCGTGAGCTCTCTTCAATCTAAGTCTCTCAAGAGATTCAG 660  
DB 625 GCATCTCTATGATGATCGTGAGCTCTCTTCAATCTAAGTCTCTCAAGAGATTCAG 684  
QY 661 GAATACAGAGATGGCGTCCCAAAATTCACAGCTGTATTCAGGTGAGAGATGCAGAA 720  
DB 685 GAATACAGAGATGGCGTCCCAAAATTCACAGCTGTATTCAGGTGAGAGATGCAGAA 744  
QY 721 ATGATGCAAGATAGGCTTCTCAATGAGATCAAAATTCATCAAGTCAAGATGGGGCA 780  
DB 745 ATGATGCAAGATAGGCTTCTCAATGAGATCAAAATTCATCAAGTCAAGATGGGGCA 804  
QY 781 AAGACTACCCAGATGATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGGCAATAT 840  
DB 805 AAGACTACCCAGATGATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGGCAATAT 864  
QY 841 CCAAGACAGGTTGATCTGATCACTGATGATCACTGATGATCACTGATGATCACTGATG 900  
DB 865 CCAAGACAGGTTGATCTGATCACTGATGATCACTGATGATCACTGATGATCACTGATG 924  
QY 901 ATGATATAGCGGTGAGAGCTTATATCAATGAGAGACATCTCACTTATTAAGATCTT 960  
DB 925 ATGATATAGCGGTGAGAGCTTATATCAATGAGAGACATCTCACTTATTAAGATCTT 984  
QY 961 GGGCTGCTCCAAAGAGAGACTGCGGCTGATGATGATGATGATGATGATGATGATGATG 1020  
DB 985 GGGCTGCTCCAAAGAGAGACTGCGGCTGATGATGATGATGATGATGATGATGATGATG 1044  
QY 1021 GTTGTGCTTCCAGATATATCAATGATGATGATGATGATGATGATGATGATGATGATG 1080  
DB 1045 GTTGTGCTTCCAGATATATCAATGATGATGATGATGATGATGATGATGATGATGATG 1104  
QY 1081 ATGAGATCTAG 1140  
DB 1105 ATGAGATCTAG 1164  
QY 1141 GCCAGGCGCATGAG 1200

[illegible]

Contact: nyc.mc@cshgri.nth.gov  
 Ahlter, N., Ayele, K., Beckstrom-Sternberg, S. M., Benjamin, B.,  
 Blakesley, R. W., Bouffard, G. G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N. L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,  
 Maduro, Q. L., Mastello, C., Mastreri, B., Mastrian, S. D., McCloskey, J. C.,  
 McDowell, J., Pearson, R., Stankitop, S., Thomas, P. J., Touchman, J. W.,  
 Tsurugan, C., Vogt, J. L., Walker, M. A., Wetherby, K. D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E. D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/INL at: <http://image.lnl.gov>  
 Series: IRAX Plate: 81 Row: 9 Column: 18  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9055233.

FEATURES  
 source  
 1..1796  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="MGC:46857 IMAG:4989536"  
 /tissue\_type="Colon, normal, 5 month old male mouse."  
 /clone\_id="NCI CGAP \_Co24"  
 /lab\_host="DH10B"  
 /note="vector: pCMV\_SPORT6"  
 1..1796  
 /gene="Pscp-pending"  
 /note="synonyms: Hls2, 1190003D2Rik"  
 /db\_xref="LOCUSID:54381"  
 /db\_xref="MGI:1889205"  
 63..1475  
 /codon\_start=1  
 /product="plasma glutamate carboxypeptidase"  
 /protein\_id="AAH37067.1"  
 /db\_xref="GI:22477498"  
 /db\_xref="LOCUSID:54381"  
 /translat="MRSLLPFLFVHLALSGKAVPKNGVSORTFREIEKIANVEDYAKALINLVYGVKONRSYERLGLVDYVGRLSKSLKVALIOIYMNLOODGLGVHLEQRIPIHMERGESAYMLERPIHKMAVLGLGSSIGTSPGIGTAVLVNLSFDLQRLASEKRGKIVYNQPYETGVQVQVRYGVAEAAVGVASILQSVASPIYSPTGICIKYDQGPRIPLACTIVEDADMMSSMAKRGKIVLHLEMGATPYDPTSFNVAITGSKYPEEVVLVSGHDSMDVGGGALDDGGGAFISWALSIVYDLGLRPKRTLVLTAEEOGGIGASQYVELLHKANISKYSLVMEADSGFLPTGGLQFDSDKARAIMEKVNMLQPLNTKVPNSGEQTDINFIQAGVGVGASLRDLYKVPFFHSHSDTMTVMDPKQNVAAAWAAVYVYVVAADMDKPRSS"

BASE COUNT  
 526 a 372 c 418 g 480 t

ORIGIN  
 Query Match 78.9%; Score 116.8; DB 10; Length 1796;  
 Best Local Similarity 87.9%; Pred. No. 3,1e-308;  
 Matches 1217, Conservative 0; Mismatches 167; Indels 0; Gaps 0;

33 TGTTCACCTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGATGGCATCTCTAAGAG 92  
 Db TGTTCACCTTTATCCCTGTGCTCTGGGCTGTGTTAAAGCTGTATTCAGATGGTCTTCTCAGCG 148  
 89 TGTTCACCTTTATCCCTGTGCTCTGGGCTGTGTTAAAGCTGTATTCAGATGGTCTTCTCAGCG 148  
 93 GACTTTGAAGAATTAAGAAGAAATAGCAGCTGTGGAGATCTCTGTAAGAATCAT 152  
 Db AACCTTTGAGAAATTAAGAAGAAATAGCAGCTGTGTAAGAATGTGCTTAAGAATTAAT 208  
 149 AACCTTTGAGAAATTAAGAAGAAATAGCAGCTGTGTAAGAATGTGCTTAAGAATTAAT 208  
 153 CAACCTAGCTGTTATGTTAAAGCCCAAGACAGATCTTATGAGCGATTTGGCACTTCTG 212  
 Db CAACCTAGCTGTTATGTTAAAGCCCAAGACAGATCTTATGAGCGATTTGGCACTTCTAGT 268  
 209 CAACCTAGCTGTTATGTTAAAGCCCAAGACAGATCTTATGAGCGATTTGGCACTTCTAGT 268  
 213 TGATATCTTTGAGCCAGACTGAGTGCCTCCAGAACCTTGAAGAAAAGCCATCAATAT 272  
 Db TGATATCTTTGAGCCAGACTGAGTGCCTCCAGAACCTTGAAGAAAAGCCATCAATAT 328  
 269 TGATATCTTTGAGCCAGACTGAGTGCCTCCAGAACCTTGAAGAAAAGCCATCAATAT 328  
 273 GTACCAAAACCTGAGAGAAGATGGGCGTGAAGAAAGTCACTGGAGCAGTGAAGAAATACC 332  
 329 GTACCAAAACCTGAGAGAAGATGGGCGTGAAGAAAGTCACTGGAGCAGTGAAGAAATACC 388

```

Qy 333 CCATGGGAGAGGGGAGAGAGATCAGCTGTGATCTGGAGCCAGAAATTCATTAAGATAGC 392
Db 389 CCACTGGGAGAGGGGAGAGAGATCTGCAGTATGATCTTGAGCTCGAAATTCAGAAATGAGC 448
Qy 393 CATCTGGGCTCTTGGGAGAGAGATTTGGGACTCTCCAGAGAGGCTTACAGAGAGATTTCT 452
Db 449 TATTTAGGCTCTTGGGAGAGAGATTTGGGACTCTCCAGAGAGGCTTACAGAGAGATTTCT 508
Qy 453 GGTGTGACCTCTTTCGATGAACTGACAGAGAGGGGCTCCAGAACCAAGAGAGAGATTTGT 512
Db 509 GGTGTGACCTCTTTCGATGAACTGAAAGAGAGCTCCAGAGAGAGAGAGAGATTTGT 568
Qy 513 TGTTTATACCAACCTTACATCACTACTCAAGAGAGGCTGCAATACGAGAGAGAGAGATTT 572
Db 569 TGTTTATACCAACCTTACATCACTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 628
Qy 573 GGTGTGAGCTGCGCAGAGAGTGGGGCTTTGGGACTCTCTCAATTCGATTCGCTGCTCTTCTC 632
Db 629 TGTGTGAGCTGCGCAGAGAGTGGGGCTTTGGGACTCTCTCAATTCGATTCGATTCGCTTCT 688
Qy 633 CATCTAGCTCTCAACAGGATTTCAAGAAATACAGAGAGGCTGCGCAGAGAGAGATTTCAAC 692
Db 689 CATCTAGCTCTCAACAGGATTTCAAGAAATACAGAGAGGCTGCGCAGAGAGAGATTTCAAC 748
Qy 693 AGCCTGATTTACAGGAGAGAGAGAGAGAGATTTCAAGAAATGAGTCTCTCAAGAGATTTCA 752
Db 749 AGCCTGATTTACAGGAGAGAGAGAGAGAGATTTCTCGAATGAGTCTCTCTGAGAGAGAG 808
Qy 753 AATTGTCATTCAGCTAAAGATGGGGGCAAGAGAGCTTACCCAGAGATCTGATTCCTTCAAC 812
Db 809 AATTGTCATTCAGCTAAAGATGGGGGCAAGAGAGCTTACCCAGAGATCTGATTCCTTCAATAC 868
Qy 813 TGTAGAGAGATTCAGCTGGGAGAGAGATTTCAAGAGAGGTTGATGATGATGATGATGATCT 872
Db 869 TGTAGAGAGATTCAGCTGGGAGAGAGATTTCAAGAGAGGTTGATGATGATGATGATGATCT 928
Qy 873 GAGAGAGAGAGATTTGGGAGAGAGAGAGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAG 932
Db 929 GAGAGAGAGAGATTTGGGAGAGAGAGAGAGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAG 988
Qy 933 GAGAGAGAGAGATTTGAGAGAGAGAGAGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAG 992
Db 989 GAGAGAGAGAGATTTGAGAGAGAGAGAGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAG 1048
Qy 993 GCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1052
Db 1049 GCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1108
Qy 1053 GGTAAATATTTTCAACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1112
Db 1109 GGTAAATATTTTCAACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1168
Qy 1113 TGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1172
Db 1169 TGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1228
Qy 1173 GGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1232
Db 1229 TGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1288
Qy 1233 TTTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1292
Db 1289 TTTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1348
Qy 1293 CTTCATCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1352
Db 1349 TTTTCATCACTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1408
Qy 1353 TGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1412
Db 1409 TGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1468
Qy 1413 GTCC 1416

```

```

Db 1469 GTCC 1472
RESULT 10
AF107835
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..1596
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="LCH-M1"
/tissue_type="liver"
184..1596
/note="secretory protein"
/codon_start=1
/product="aminopeptidase"
/db_xref="GI:5442032"
/translation="MRSIFPLFTVHLAAGSKAVFNGVDSORPREKEIANVRY
AKKINLNAVYKQNSRYELGLVDIVGRSLSGSKRLBKAIQIMQNLQDDGLNRY
LEQVRIPIHMERGESAVMEPRIRHNAIILAGSISIGPPGGLITAEVLVAVSPELQNR
ASBARGLIYVNOPTGYEKTQYRQVGAFAKAVGAVLQSVASFYSFPTGIG
KYDDGPKIPTACTIYEDAMSRMSRGRKIVIHLEMGAKTVPDTSFNTVAITGS
MYEEVAVLVSGHLSMDVGGALDDGGARISMEALSLVLDGLRPRKTRLVIMTAE
EOGGIAGASOYVELHXANISKYSLVMEADSGTLPGLQFTGSDKARAIKENVNLIQ
LANTKFSNBEGRDINFWIAGVPGVSLRDLVYKFFHSHSDTMTVMDPKQMNVA
AVMAVVAAYVVDMDMLPRS"
BASE COUNT 415 a 357 c 421 g 403 t
ORIGIN
Query Match 78.8%; Score 1115.2; DB 10; Length 1596;
Best Local Similarity 87.9%; Pred. No. 8-be-308;
Matches 1216; Conservative 0; Mismatches 168; Indels 0; Gaps 0;
Qy 33 TGTTCACCTTTTATCCCTGTGCTCTGGGAGAGATTAATGCAAGATGCAATCTTAAGAG 92
Db 210 TGTTCACCTTTTATCCCTGTGCTCTGGGAGAGATTAATGCAAGATGCAATCTTAAGAG 269
Qy 93 GACTTTTGAAGAAATTAAGAGAGATAGCCAGCTGTGAGATGTTGCTTAAGCAATCAT 152
Db 270 AACATTTGAGAAATTAAGAGAGATAGCCAGCTGTGAGATGTTGCTTAAGCAATCAT 329
Qy 153 CAACCTAGCTTTTATAGTAAAGCCGAGAGAGAGATCTTAAGAGAGATGTTGCTTCTGCT 212
Db 330 CAACCTAGCTTTTATAGTAAAGCCGAGAGAGAGATCTTAAGAGAGATGTTGCTTCTGCT 389
Qy 213 TGAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 272
Db 390 TGAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 449
Qy 273 GATCAAAACCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 332
Db 450 GATCAAAACCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 509

```

```

Qy 333 CCACGTGGAGAGGGAGAGAGATCAGTGTGATGCTGGAGCCAGAAATTCATTAAGATAGC 392
Db 510 CCACTGGAGAGGGAGAGAGATCTGAGTATGTTGAGCTGGAATTCAGAGATGGC 569
Qy 393 CATCTGGGTCTTGGACAGACATTTGGAGCTCTCCAGAGAGCATTTACAGAGAAATTTCT 452
Db 570 TATCTGAGTCTTGGACAGACATTTGGAGCTCTCCAGAGAGCATTCACAGAGAAATGCT 629
Qy 453 GGTGTGACCTCTTTCAGTCAAGTGCAGAGAGGGCTCAGAGAGAGAGAGAGATTTGT 512
Db 630 GGTGTGACCTCTTTCAGTCAAGTGCAGAGAGGGCTCAGAGAGAGAGAGATTTGT 689
Qy 513 TGTATTAAACCACTTACATCACTACTCAAGAGCGGTGCAATACCGAGCGAGGGGC 572
Db 690 TGTATTAAACCACTTACATCACTACTCAAGAGCGGTGCAATACCGAGCGAGGGGC 749
Qy 573 GGTGAGAGCTCCAGAGGTGGGGCTTTGGCATCTCTCATTTGATCCGTGGCTCTTCTC 632
Db 750 TGTGGAAGCTCCAGAGGTGGGGCTTTGGCATCTCTCATTTGATCCGTGGCTCTTCTC 809
Qy 633 CATCTACAGTCTCAACAGGATTTCAAGATACAGAGATGGCGCCAAATTCACAC 692
Db 810 CATCTACAGTCTCAACAGGATTTCAAGATACAGAGATGGCGCCAAATTCACAC 869
Qy 693 AGCTGTATTACGGTGAAGATGCAAGATATGTCAGAAATGGCTTCTCAATGGATCAA 752
Db 870 AGCTGTATTACGGTGAAGATGCAAGATATGTCAGAAATGGCTTCTCAATGGATCAA 929
Qy 753 AATTGTCACTCACTTAAGATGGGGGCAAGACCTACCCAGATCTGATTCCTTCAACAC 812
Db 930 AATTGTCACTCACTTAAGATGGGGGCAAGACCTACCCAGATCTGATTCCTTCAACAC 989
Qy 813 TGTGAGAGAGATCACTGGAGCAATATTCAGAAACAGTGTGATCTGTCAGTGGACATCT 872
Db 990 TGTGAGAGAGATCACTGGAGCAATATTCAGAAACAGTGTGATCTGTCAGTGGACATCT 1049
Qy 873 GAGACAGTGGAGATTTGGGCGAGGTGCCATGGATGATGGCGTGGAGCTTTATATCATG 932
Db 1050 GAGACAGTGGAGATTTGGGCGAGGTGCCATGGATGATGGCGTGGAGCTTTATATCATG 1109
Qy 933 GGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTCTGGCGCTGCT 992
Db 1110 GGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTCTGGCGCTGCT 1169
Qy 993 GCTGTGAGTGCAGAGAAAGAGAGTGGTGGCTCTCCAGTATTTACATTACACAA 1052
Db 1170 GCTGTGAGTGCAGAGAAAGAGAGTGGTGGCTCTCCAGTATTTACATTACACAA 1229
Qy 1053 GGTAAATATTTTCACTACAGTCTGTGATGAGTCTGACGAGAGACCTTCTTACCAC 1112
Db 1230 GGTAAATATTTTCACTACAGTCTGTGATGAGTCTGACGAGAGACCTTCTTACCAC 1289
Qy 1113 TGGGCTGCAATTCACGTGCAAGTGAAGAGCGAGCCATCATGAGAGAGGTTATGAGCTT 1172
Db 1290 TGGGCTGCAATTCACGTGCAAGTGAAGAGCGAGCCATCATGAGAGAGGTTATGAGCTT 1349
Qy 1173 GCTGAGCCCTCATATATCATCTCAGGTCCTGAGCCATGAGAGAGAGAGAGACATCAATT 1232
Db 1350 TCTGCAACCCCTCATATATCATCTCAGGTCCTTATGATGAGAGAGAGAGACATCAATT 1409
Qy 1233 TTTGATCCAGAGTGGAGTGGCTGGAGCAGTCTACTTATGATCACTTATCAAGTATTTCTT 1292
Db 1410 TTTGATCCAGAGTGGAGTGGCTGGAGCAGTCTACTTATGATCACTTATCAAGTATTTCTT 1469
Qy 1293 CTTTCATCACTCCACGAGAGACCATGATCTGTCAATGATCCAAAGAGATGATTTGC 1352
Db 1470 TTTTCATCACTCCACGAGAGACCATGATCTGTCAATGATCCAAAGAGATGATTTGC 1529
Qy 1353 TGTCTGCTGTTGGGTGTTGTTTCTTAAGTGTGTCAGACATGAGAGAAATGCTGCCATG 1412
Db 1530 TGTCTGCTGTTGGGTGTTGTTTCTTAAGTGTGTCAGACATGAGAGAAATGCTGCCATG 1589
Qy 1413 GTCC 1416

```

```

Db 1590 GTCC 1593
RESULT 11
AF131077 1726 bp mRNA linear ROD 03-JAN-2003
LOCUS Rattus norvegicus liver annexin-like protein (LAL) mRNA, complete
DEFINITION cds.
ACCESSION AF131077
VERSION AF131077.1 GI:7108712
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1726)
Della Fazio, M.A., Piobbico, D., Bartoli, D., Castell, M.,
Brancorsini, S., Viola Magni, M., and Servillo, G.
lal-1: a differentially expressed novel gene during proliferation
in liver regeneration and in hepatoma cells
Genes Cells 7 (11), 1183-1190 (2002)
JOURNAL MEDLINE 22278398
PUBMED 12390252
AUTHORS Servillo, G., Della Fazio, M.A., Piobbico, D., Bartoli, D.,
Castell, M., Brancorsini, S. and Viola Magni, M.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-1999) Institute of General Pathology, University
of Perugia, Policlinico Monteluce, Perugia 06100, Italy
FEATURES
location/Qualifiers
source 1..1726
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
1..1726
/gene="LAL"
63..1481
/gene="LAL"
/note="Involved in liver regeneration; contains unique
annexin domain at the N-terminus"
/codon_start=1
/product="liver annexin-like protein"
/protein_id="AAP6518.1"
/db_xref="GI:7108713"
/translaction="MRFLFLVAVAVHLSGSKAIYKSGVSGRTQRIKEIENYE
VHAEVYIPIHMERGESBVAWVPRILHAIILIGSGSIGTPEBGLTAEVLVASFVELO
RRASERAGKIVVNOPPYTDGKTVQYERGAVALAKYASILRSVASFSISPTG
HOGVODGVPRIPIACITIEDAEMMSRTRASGDKIVILHAKGAKTYPTDSFNTVAEIT
GSKYPERVAVVSHLDSMDVQCALDGGCAFLSWELSLVKGILRPRKTRILVMT
AESRGVVASQVYELHAKANISKYSLVNBAUSGTFPLPTGLPTGSDKRAIMKEVMSIL
OPLNITKVFNDAGSTIDINFWIOGVDPASLRDLKYFFHSHGDTMTAMDPRKMNW
AAAVMAVAVVAVVADMEMLPRS"
BASE COUNT 474 a 385 c 428 g 439 t
ORIGIN
Query Match 78.2%; Score 1107.2; DB 10; Length 1726;
Best local similarity 86.4%; Pred. No. 1,7e-305;
Matches 1223; Conservative 0; Mismatches 193; Indels 0; Gaps 0;
Qy 1 ATGAAATTCCTTATCTTGGATTTTGGGCTTACCTTTATCCCTGTGCTGGG 60
Db 63 ATGAGGTTCCTTTCTTCTGTTGCTGTTGTTTCACTTTTCTCTGGGCTTGGGA 122
Qy 61 AAAGCTATGCAAGATGCGATCTTAAGAGAGCTTTTGAAGAAATAAAGAAATA 120
Db 123 AAAGCTATATCAAGATGCGTCTTCAAGCAACATTTCAAGAAATAAAGAAATA 182
Qy 121 GCGAGCTGTGAGATGTTGCTAAAGCATCATCACTAGCTGTTATGTTAAAGCCAG 180
Db 183 GCCAACTATGAAAGATGTTGCTAAAGCAATATCAACCTGCTGTTATGGAATAACAC 242

```

QY 181 AACGATCTATGAGGATTTGGCACTTCTGTGATATCTGTGGAGCCAGAGTGAAGG 240  
 Db 243 AACCGTCTATGAGGCTTTGGAGCTTCTGTGATATCTGTGGAGCCAGAGTGAAGG 302  
 QY 241 TCCAGAACCTAGAAAAAGCATCCAAATTAATGATCCAAAACTGCAGCAAGATGGGCTG 300  
 Db 303 TCTAAGAACCTAGAAAAAGCATCCAAATTAATGATCCAAAACTGCAGCAAGATGGGCTG 362  
 QY 301 GAGAAAGTTCACTGGAGGCAAGTGAATACCCCACTGGAGAGAGGAGAAAGATCAAGCT 360  
 Db 363 GAAAGCTTCACCTGGAGGCAAGTGAATACCCCACTGGAGAGAGGAGAAAGATCAAGCT 422  
 QY 361 GTGATGCTGAGGCAAAATTCATTAAGATGACCTCTGGGCTTTGGCAGAGCATTTGG 420  
 Db 423 GTGATGCTGAGGCTTCGAATTCACAAAGTGGCTATTTAGGCTTTGGCAGAGCATTTGG 482  
 QY 421 ACTCTCTCAGAAAGCATTAACAGCAAGATTCCTGTGTGATCCTCTTTGATGAATCTGAG 480  
 Db 483 ACTCTCTCAGAAAGCATTAACAGCAAGATTCCTGTGTGATCCTCTTTGATGAATCTGAG 542  
 QY 481 AGAAGGCTCTCAGAAAGCAAGAGGAGATTTGTTTATTAACCACTTACATCACTAC 540  
 Db 543 AGAAGGCTCTCAGAAAGCAAGAGGAGATTTGTTTATTAACCACTTACATCACTAC 602  
 QY 541 TCAAGAGCGGTGCAATCCGAAAGCGAGGCGGTGAGAGCTGCAAGGTGGGCGCTTTG 600  
 Db 603 GGGAAAGCTGTGAGTACCGAGAGCGGAGCTGTGAGAGCTGCAAGGTGGGCGCTTTG 662  
 QY 601 GCATCTCTCATTCGATCCGTGGCTCTCTTCATCTCAAGTCTCTCAACAGATTAATCAG 660  
 Db 663 GCATCTCTCATTCGATCCGTGGCTCTCTTCATCTCAAGTCTCTCAACAGATTAATCAG 722  
 QY 661 GAATACCAAGATGCGGCGGCAAAATTCCAACAGCTGTATTAAGGTGAAGATGCGAA 720  
 Db 723 GATATCAAGATGCGGCGGCAAAATTCCAACAGCTGTATTAAGGTGAAGATGCGAA 782  
 QY 721 ATGATGTCAGAAAGTGTCTTCATGAGGATCAAAATTCATCAAGCTTAAGATGGGGCA 780  
 Db 783 ATGATGTCAGAAAGTGTCTTCGAGGAGCAAAATTCATCAAGCTTAAGATGGGGCA 842  
 QY 781 AAGACTTACCAAGATGATGATCTCTTCAACATCTAGCAGAGATCACTGGAGCAAAAT 840  
 Db 843 AAGACTTACCAAGATGATGATCTCTTCAACATCTAGCAGAGATCACTGGAGCAAAAT 902  
 QY 841 CCAGAAAGATTGATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
 Db 903 CCAGAAAGATTGATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 962  
 QY 901 ATGATGATGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 Db 963 CTGATGATGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1022  
 QY 961 GGGCTGCTCCAAAGAGGATCTGCGGCTGTGCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1020  
 Db 1023 GGGCTGCTCCAAAGAGGATCTGCGGCTGTGCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 1082  
 QY 1021 GTTGTGCTCTCCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 Db 1083 GTTGTGCTCTCCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1142  
 QY 1081 ATGATGATGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
 Db 1143 ATGATGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1202  
 QY 1141 GCGAGGCTCATGAGAGGATTAATGAGCTGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1200  
 Db 1203 GCGAGGCTCATGAGAGGATTAATGAGCTGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1262  
 QY 1201 CTGAGCTCATGAGAGGATTAATGAGCTGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1260  
 Db 1263 TTTATGATGAGAGGATTAATGAGCTGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1322

QY 1261 AGTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 Db 1323 AGTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1382  
 QY 1321 ACTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
 Db 1383 ACTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1442  
 QY 1381 GTTGTGAGAGATGAG 1416  
 Db 1443 GTTGTGAGAGATGAG 1478  
 RESULT 12  
 AX400829  
 LOCUS 1778 bp DNA linear PAT 06-JUN-2002  
 DEFINITION Sequence 505 from Patent WO0210453.  
 ACCESSION AX400829  
 VERSION AX400829.1 GI:21337009  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE  
 1 Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and  
 Elashoff, M.R.  
 TITLE Molecular toxicology modeling  
 JOURNAL Patent: WO 0210453-A 505 07-FEB-2002;  
 Gene Logic, Inc. (US)  
 FEATURES  
 source  
 1..1778  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /note="EMBL/GenBank Accession No. AF097723"  
 BASE COUNT 488 a 391 c 442 g 457 t  
 ORIGIN  
 Query Match 78.2%; Score 1107.2; DB 6; Length 1778;  
 Best Local Similarity 86.4%; Pred. No. 1,7e-305;  
 Matches 1223; Conservative 0; Mismatches 193; Indels 0; Gaps 0;  
 QY 1 ATGAAATTCCTTAATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60  
 Db 121 ATGAGTTCCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 QY 61 AAGCTATATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
 Db 181 AAGCTATATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
 QY 121 GCCAGCTGTGAGAGATGCTTAAAGCAATCACTAGCTGTTATGTTAAAGCCAG 180  
 Db 241 GCCAATATGAAAGATGTTGCTTAAAGCAATATCACTGCTGTTATGAAATATCAG 300  
 QY 241 AACGATTCATGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
 Db 301 AACCGTGTATGAGCCTTTGGAGCTTCTAGTTGATATCTGTGACCCAGACTGAGTGGC 360  
 QY 301 AACCGTGTATGAGCCTTTGGAGCTTCTAGTTGATATCTGTGACCCAGACTGAGTGGC 360  
 Db 361 TCTAAGAACCTAGAAAAAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
 QY 421 GAGAAAGTTCACTGAGGCAAGTGAATACCCCACTGGAGAGAGGAGAGAAATCAAGCT 360  
 Db 421 GAAAGCTTCACCTGAGGCAAGTGAATACCCCACTGGAGAGAGGAGAGAAATCAAGCT 480  
 QY 361 GTGATGCTGAGGCAAAATTCATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 420  
 Db 481 GTGATGCTGAGGCTTCGAATTTCAAGATGATGATGATGATGATGATGATGATGATGATGAT 540  
 QY 421 ACTCTCAGAAAGCATTAACAGCAAGATTCCTGTGTGATCCTTTGATGAATCTGAG 480

```

Db 541 ACTCTCTCTGAGGATACAGCAGAGAACTACTCGTGGGCTCTTTTGTGAATCAA 600
Qy 481 AGAAGGCTCTGAGAGCAGAGAGGAGATGTTGTTTAAACCACTTACATCACTAC 540
Db 601 AGAAGGCTCTGAGAGCAGAGAGGAGATGTTGTTTAAACCACTTACATCACTAC 660
Qy 541 TCAAGAGCGGTCATACGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
Db 661 GGGAAATCTGCTAGTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
Qy 601 GCATCTCTCATTCGATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Db 721 GCATCTCTCATTCGATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Qy 661 GAATACCAAGATGCGGCTCCCAAAATTCGAGGAGGAGGAGGAGGAGGAGGAG 720
Db 781 GGATATCAAGATGCGGCTCCCAAAATTCGAGGAGGAGGAGGAGGAGGAGGAG 840
Qy 721 ATGATGTCAGAAATGCGCTCTCATTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Db 841 ATGATGTCAGAAATGCGCTCTCATTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Qy 781 AAGACCTACCGAGATCTGATTCCTTCAACACTGAGCAGAGATCACTGGAGCAATAT 840
Db 901 AAGACCTACCGAGATCTGATTCCTTCAACACTGAGCAGAGATCACTGGAGCAATAT 960
Qy 841 CCAAGACAGGTTGTAATCTGTCAGTGAATCTGCACTGGAGATGTTGGGAGGAGTCC 900
Db 961 CCAAGACAGGTTGTAATCTGTCAGTGAATCTGCACTGGAGATGTTGGGAGGAGTCC 1020
Qy 901 ATGATGTCAGAAATGCGCTCTCATTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
Db 1021 CTGATGTCAGAAATGCGCTCTCATTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Qy 961 GGGGCTCTCCAAAGAGAGACTCTGCGGCTGTCCTGAGTGCAGAGAAACAAGTGA 1020
Db 1081 GGGGCTCTCCAAAGAGAGACTCTGCGGCTGTCCTGAGTGCAGAGAAACAAGTGA 1140
Qy 1021 GTTGGTCTCTCCAAAGAGAGACTCTGCGGCTGTCCTGAGTGCAGAGAAACAAGTGA 1080
Db 1141 GTTGGTCTCTCCAAAGAGAGACTCTGCGGCTGTCCTGAGTGCAGAGAAACAAGTGA 1200
Qy 1081 ATGATGTCAGAAATGCGCTCTCATTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db 1201 ATGATGTCAGAAATGCGCTCTCATTCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Qy 1141 GCCAGGCTCATTCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Db 1261 GCCAGGCTCATTCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Qy 1201 CTGAGCCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Db 1321 TTTTATGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
Qy 1261 AGTCTACTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
Db 1381 AGTCTACTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Qy 1321 ACTGTCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
Db 1441 ACTGTCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
Qy 1381 GTTGTGCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Db 1501 GTTGTGCAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560

```

RESULT 13  
 AF097723 1778 bp mRNA linear ROD 10-NOV-1998  
 LOCUS Rattus norvegicus hematopoietic lineage switch 2 related protein  
 DEFINITION (H1s2-rp) mRNA, complete cds.

```

ACCESSION AF097723
VERSION AF097723.1 GI:3851631
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (base 1 to 1778)
AUTHORS Chen, Y. and Talmage, D.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1998) Institute of Human Nutrition, Columbia
University, 701 West 168th Street Room 5-503, New York, NY 10032,
USA
FEATURES
Source location/Qualifiers
1..1778
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Fisher"
/db_xref="taxon:10116"
/cell_line="F11"
/cell_type="fibroblast"
/dev_stage="embryo"
1..1778
/gene="H1s2-rp"
121..1539
/gene="H1s2-rp"
/codon_start=1
/product="hematopoietic lineage switch 2 related protein"
/protein_id="AAC72384.1"
/db_xref="GI:3851632"
/translat="MRFLPFLFAVAVHLFSLGSGKATYKSGVRSORFQIKESTANE
DVAKATINIAVYGRKQSRSEYERGLVDYTPGLSGSKNLEKAIQIMVLODGLN
VHLEQVRIPEHGBSAYVVRIRHIALILGJSGTPEEGITAVLVASVEIQ
RRASEARKIIVVNPPTDYGKTQYKERAVBAKGAVALIRVASFSLSPHTG
HOGYQDEVPKIPRACITIEDAEMMSRMSAGDIXIHLKMGAKTYPTDSEFNTVAET
GSKYPERDLVSGHLSMDVGGALDGGAFISWEALSLVKDLGLPKRTLRVLT
ABRGGVAGSIOYIEHKNANISKYSLVWEADSGFLPTGLQFTSGDKARAIKMEVMSL
QPNATYVPRNDAEGTDINPWICQVAGRPASLRDDLKXKFFPHSHGDTMTMDPRQNV
AAAVMAVAIVVADMEMLPRS"
BASE COUNT 488 a 391 c 442 g 457 t
ORIGIN
Query Match 78.2%; Score 1107.2; DB 10; Length 1778;
Best Local Similarity 86.4%; Pred. No. 1.7e-305;
Matches 1223; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
1 ATGAATTCCTTATCTTGCAATTTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60
121 ATGAGGTTCTCTTCTTCTGTTGCTGTTGTCACCTTTTATCCTGCTGAGG 180
61 AAAGCTATGCAAGAAATGCGATCTTAAGAGACTTTTGAAGAAATMAAAGAAATA 120
181 AAAGCTATGCAAGAAATGCGATCTTAAGAGACTTTTGAAGAAATMAAAGAAATA 240
121 GCCAGCTGAGAGATGTTGCTAAAGCAATCAACCTAGCTGTTATGGAAGCCGAG 180
241 GCCACTATGAGAGATGTTGCTAAAGCAATCAACCTAGCTGTTATGGAAGATACAG 300
181 AACAGATCTATGAGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
301 AACGCGTGTATGAGCGTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
241 TCCTAAGAACTTGAAGAAAGGATCAATCATGTATCAAAAACCTGCAACAAGAGGCTG 420
301 GAGAAAGTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
421 GAAAACGTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
361 GTGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420

```

```

Db 481 GTGATGCTGCTGCTCAATTCACAGTTGGCTATTTTAGCCCTTGGCGGACGATTGG 540
Qy 421 ACTCTCCAGAGGCAATTACAGAGATTCTGGTGGTGAACCTTTTCATGAATGACGAG 480
Db 541 ACTCTCTCAGAGGATACAGAGGAGTACTCGTGGTGGCTTTTGTGTAACCTTCAA 600
Qy 481 AGAAGGCTTCAGAGGAGGAGGAGATTTGTTTATACCAACCTTACATCACTAC 540
Db 601 AGAAGGCTTCAGAGGAGGAGGAGATTTGTTTATACCAACCTTACATCACTAC 660
Qy 541 TCAAGAGCGGTGCAATCCGACCGGAGGCGGTGAGAGCTGCGCAAGGTGGGGCTTGG 600
Db 661 GGGAAAATCTGCAATCCGAGGAGGCGGAGCTGTGAGAGCTGCGCAAGGTGGGGCGGTG 720
Qy 601 GCATCTCTCATGTCGATCCGAGGCTCTCTCCATCTACAGTCTCTCAACAGGATTCAG 660
Db 721 GCATCTCTCATGTCGATCCGAGGCTCTCTCCATCTACAGTCTCTCAACAGGATTCAG 780
Qy 661 GAATACCAAGATGGCGTCCGCAAAATTCACACAGCTGTATTCAGGTGAGAGATGAGAA 720
Db 781 GATATCAAGATGGTGGTCCCAAGATTCACACAGCTGTATTCAGATGAGAGATGAGAA 840
Qy 721 ATGATGCAAGATGGTCTCTCATGAGATCAAAATTTGCTATCAGCTAAAGATGGGGCA 780
Db 841 ATGATGCTGAAATGGCTCTCTCGTGGGAGCAAAATTTGCTATCTGAAATGGGAGCA 900
Qy 781 AAGACCTACCAAGATGATGATCTTCAACACCTGACAGAGATCACTGGGAGGCAATAT 840
Db 901 AAGACCTATCAAGATGATGATCTTCAACACCTGACAGAGATCACTGGGAGGCAATAT 960
Qy 841 CCAGAACAGGTTGATCTGTCAGTGAATCTGTCAGCTGGAGTGTGGGCAAGGCTGCC 900
Db 961 CCAGAGAGATGTTGCTGTCAGTGAATCTGTCAGCTGGAGTGTGGGCAAGGCTGCC 1020
Qy 901 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Db 1021 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Qy 961 GGGCTGCTGCAAAAGAGAGATCTGGGCTGGTCTGGAATGAGAGAGAGAGAGAGAG 1020
Db 1081 GGGCTGCTGCAAAAGAGAGATCTGGGCTGGTCTGGAATGAGAGAGAGAGAGAGAG 1140
Qy 1021 GTTGTGCTCTTCAATTTACATTTACATTTACATTTACATTTACATTTACATTTAC 1080
Db 1141 GTTGTGCTCTTCAATTTACATTTACATTTACATTTACATTTACATTTACATTTAC 1200
Qy 1081 ATGAGATCTGACGAGAGACCTTTTACCACTGGGCTGCAATCTGTCAGTGAAG 1140
Db 1201 ATGAGGCTGACTCAGAGACCTTTTACCACTGGGCTGCAATCTGTCAGTGAAG 1260
Qy 1141 GCCAGGCTCATGATGAGAGGTTATAGCTGCTGAGCCCTCATATTCACATCTCAG 1200
Db 1261 GCCAGGCTCATGATGAGAGGTTATAGCTGCTGAGCCCTCATATTCACATCTCAG 1320
Qy 1201 CTGAGCCATGAGAGAGGAGAGACATCACTTTTGAATCCAGCTGAGTGGCTGAGCC 1260
Db 1321 TTTAATGATGAGAGAGGAGATCACTTTCGATTCAGAGTGGAGTGGCTGAGCC 1380
Qy 1261 AGCTTACTGATGATCAATTAACAATATTTCTTCTTCATCACTCCACGAGAGACCATG 1320
Db 1381 AGCTTACTGATGATCAATTAACAATATTTCTTCTTCATCACTCCACGAGAGACCATG 1440
Qy 1321 ACTGTCATGATCCAAAGCAGATGAAATGTCGCTGCTGTTGGGCTGTTGTTCTTAT 1380
Db 1441 ACTGTCATGATCCAAAGCAGATGAAATGTCGCTGCTGTTGGGCTGTTGTTCTTAT 1500
Qy 1381 GTTGTGCTGAGATGAGAGAAATGCTGCTGAGTCC 1416
Db 1501 GTTGTGCTGAGATGAGAGAAATGCTGCTGAGTCC 1536

```

RESULT 14  
AF009513

```

LOCUS AF009513 1716 bp mRNA linear ROD 01-JUN-1998
DEFINITION Mus musculus hematopoietic lineage switch 2 (HLS2) mRNA, complete
cde.
ACCESSION AF009513
VERSION AF009513.1 GI:3169728
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Williams,J.H., Chan,C.-Y. and Klinken,S.P.
TITL Hematopoietic lineage switch 2 (HLS2), a novel mRNA species induced
during an erythroid to myeloid lineage switch
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1716)
AUTHORS Williams,J.H., Chan,C.-Y. and Klinken,S.P.
TITL Direct Submission
JOURNAL Submitted (20-JUN-1997) Biochemistry Department (Laboratory of
Cancer Medicine), University of Western Australia, Royal Perth
Hospital, Wellington Street, Perth, WA 6001, Australia
FEATURES
source
1..1716
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..1716
/gene="HLS2"
97..1398
/gene="HLS2"
/product="hematopoietic lineage switch 2"
/codon_start=1
/protein_id="AA17945.1"
/db_xref="GI:3169729"
/translaction="MRSLEFLPLVHLALSGKAVKNGVORTFREIKEIANYEDV
AKAIIINLAVYKXONSVERGLVDTVPKRSKNEKALQIOMQNLQODLENVH
LEQVRI.PHMERGESVAMLEPRHMAIITGLSGISCTPPGTLAEVLVAASFELGR
ASEBARKIIVNOPYTVERKVOYRVQAGVAKVAHVAKVAHVAKVAHVAKVAHVAKVAH
KXODGPKI.PTACITVEDAEKMSRMSRSCNKIVIHENAKTYPPDPSNTVAETIS
MPREDVGGALDDGGAFISWEALSLVKOLGRPKETLVLMTAEQGTGASQYE
LHRANYSKSLVNSADSGFLPTGLQFTSDKARAIKEMVNLLOPLNTKVFNSNEG
TDINFWIQGVNSASLRDLDLYKFFPHSHGDTMSWIOSR"
BASE COUNT 476 a 370 c 414 g 456 t
ORIGIN
Query Match 72.8%; Score 1030.8; DB 10; Length 1716;
Best Local Similarity 85.6%; Pred. No. 1.3e-283;
Matches 1185; Conservative 0; Mismatches 162; Indels 37; Gaps 2;
Qy 33 TGTTCACCTTTTATCCCTGCTGCTGAGAAAGTATATGCAAGATGCACTCTTAAG 92
Db 123 TGTTCACCTTTTATCCCTGCTGCTGAGAAAGTATATGCAAGATGCTTCTACGC 162
Qy 93 GACTTTGAAGAATATAAGAAATAGCCAGCTGTGAGAGTGTCTAAAGCAATCAT 152
Db 183 AACATTTCCGAATAATAAGAAATAGCCACTATGAAGATGTCTTAAAGCAATTA 242
Qy 153 CAACCTAGCTGTTTATGTTAAGCCAGAAACAGATCCATATGACGATGGCACTTCTGT 212
Db 243 CAACCTAGCTGTTTATGTTAAGTTAATACAGAACCCGCTCTATGAGCGTTTGGAGCTTCTAGT 302
Qy 213 TGAATAGCTGTTGAGCCAGACTGAGTGGCTCCAGAAAGCTTGAAGAAAGCCATCCAAATTA 272
Db 303 TGAATAGCTGTTGAGCCAGACTGAGTGGCTCTTAAAGAAAGCTTGAAGAAAGCTTCAAAATCAT 362
Qy 273 GTACCAAAACCTGTCAGCAAGATGGGCTGAGAAAGTTCACTCTGAGCAGTGAATACC 332
Db 363 GTACCAAAACCTGTCAGCAAGATGGGCTGAGAAATGTTCACCTGAGCAGTGAATACC 422
Qy 333 CCACCTGGAGAGGGGAGAGAGATCACTGTGATGCTGAGCCAAAGATTCATTAAGATAGC 392
Db 423 CCACCTGGAGAGGGGAGAGAGATCTGACGATGATGCTGAGCCCTGCAATTCACAGATAGGC 482

```



```

OY 393 CATCTGGGCTTTGGCAGACGATTTGGAGCTCTCCAGAAAGGCTATTACAGAGAGTTCT 452
DB 483 TATCTAGGTTCTTGGCAGACGATTTGGAGCTCTCCAGAAAGGCTATTACAGAGAGTTCT 542
OY 453 GGTGGTGAACCTTTTGCATGAACTGACAGAAAGGGCTCAGAAAGGAGGGAATTTGT 512
DB 543 GGTGGTGAACCTTTTGCATGAACTCAGAAAGGAGCATCAGAAAGGAGGGAATTTGT 602
OY 513 TGTTTTAAACCACTTTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 572
DB 603 TGTTTTAAACCACTTTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 662
OY 573 GGTGAAGCTCCCAAGTGGGGGCTTTGGAGCTCTCTCACTTGCATCCGTCCTCTCTC 632
DB 663 TGTGAAGCTCCCAAGTGGGGGCTTTGGAGCTCTCTCACTTGCATCCGTCCTCTCTC 722
OY 633 CATCTACAGTCTCTCAACAGGATTTCAAGAAATCCAGATGGCGTCCCAAAATTTCCAA 692
DB 723 CATCTACAGTCTCTCAACAGGATTTCAAGAAATCCAGATGGCGTCCCAAAATTTCCAA 782
OY 693 AGCTGTATTTACGGTGAAGATGCAAGAAATGATGTCAGAAATGCTTTCTCATGGGATCA 752
DB 783 AGCTGTATTTACGGTGAAGATGCAAGAAATGATGTCAGAAATGCTTTCTCATGGGATCA 842
OY 753 AATTGTCACTAGCTAAAGATGGGGGCAAAAGCTAACCCAGATCTGATTCCTTCAACAC 812
DB 843 AATTGTCACTAGCTAAAGATGGGGGCAAAAGCTAACCCAGATCTGATTCCTTCAACAC 902
OY 813 TGTAGCAGAGATCACTGGGAGCAAAATATCCAGAACAGTTGATGTCAGTGCATGACATCT 872
DB 903 TGTAGCAGAGATCACTGGGAGCAAAATATCCAGAACAGTTGATGTCAGTGCATGACATCT 935
OY 873 GAGACGCTGGATGTTGGGAGGCTGCAAGATGATGAGCGTGGAGGCTTTATATCATG 932
DB 936 -----GAGATGTTGGGAGGCTGCAAGATGATGAGCGTGGAGGCTTTATATCATG 986
OY 933 GAAAGCACTCTCACTTATTAAGATCTTGGGCTGGCTGCTCAAAAGAGACTTGGCGCTGT 992
DB 987 GAAAGCACTCTCACTTATTAAGATCTTGGGCTGGCTGCTCAAAAGAGACTTGGCGCTGT 1046
OY 993 GCTCTGACCTGACAGAAAGCAAGGATGAGTGGTCCCTTCCAGTATTAATCAACAA 1052
DB 1047 GCTCTGACCTGACAGAAAGCAAGGATGAGTGGTCCCTTCCAGTATTAATCAACAA 1106
OY 1053 GGTAAATTTTCCAACTACAGTCTGATGATGAGTCTGACGAGAACTTTTATACCAAC 1112
DB 1107 GGTAAATTTTCCAACTACAGTCTGATGATGAGTCTGACGAGAACTTTTATACCAAC 1166
OY 1113 TGGGCTGCAATTTCACTGGCAGTGAAGAAAGCCAGGCTCATGAGAGAGGTTATGAGCT 1172
DB 1167 TGGGCTGCAATTTCACTGGCAGTGAAGAAAGCCAGGCTCATGAGAGAGGTTATGAGCT 1226
OY 1173 GCTGAGCCCTTCAATTTCACTCAGTCTGAGCCATGAGAGAGGAGACATCAACTT 1232
DB 1227 TCTGAGCCCTTCAATTTCACTCAGTCTGAGCCATGAGAGAGGAGACATCAACTT 1286
OY 1233 TTTGATCCAGTGGAGTGGCTGAGAGCCAGTCTATTGATGACTTATTAAGATTTCTT 1292
DB 1287 TTTGATCCAGTGGAGTGGCTGAGAGCCAGTCTATTGATGACTTATTAAGATTTCTT 1346
OY 1293 CTTTCATCACTCCCAAGGAGACACCATGATGTCATGATCCAAAGAGAGATGTTGCT 1352
DB 1347 TTTTCATCACTCCCAAGGAGACACCATGATGTCATGATCCAAAGAGAGATGTTGCT 1405
OY 1353 TGTGCTGTTTGGGCTGTTGTTTCTTATGTTGTTGACAGACATGAAAGATGCTGCTAG 1412
DB 1406 TGTGCTGTTTGGGCTGTTGTTTCTTATGTTGTTGACAGACATGAAAGATGCTGCTAG 1465
OY 1413 GTCC 1416
DB 1466 GTCC 1469

```

```

RESULT 15
BD059610 609 bp DNA linear PAT 27-AUG-2002
DEFINITION Secreted expressed sequence tags (ESTs).
ACCESSION BD059610
VERSION BD059610.1 GI:22605216
KEYWORDS JP 2001519666-A/1465.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE Jacobs,K., McCoy,J.M., Lavallie,E.R., Racie,L.A., Werberg,D.,
Trecay,M., Spaulding,V. and Agostino M.J.
Secreted expressed sequence tags (ESTs)
Patent: JP 2001519666-A 1465 23-OCT-2001;
GENETICS INSTITUTE INC
PN JP 2001519666-A/1465
PD 23-OCT-2001
PR 10-APR-1998 JP 1998543068
PI 10-APR-1997 US 08/835913
PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI
DAVID MERBERG,
PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC
C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:
Double:
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source 1..609
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
BASE COUNT 190 a 127 c 148 g 144 t
ORIGIN
Query Match 36.8%; Score 520.6; DB 6; Length 609;
Best Local Similarity 99.1%; Pred. No. 1,be-137;
Matches 534; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
OY 1 ATGAATTCCTTATCTTGGATTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 60
DB 71 ATGAATTCCTTATCTTGGATTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 130
OY 61 AAAGCTATATGCAAGATGGCATCTTAAAGAGACTTTGAAGAAATTAAGAAATA 120
DB 131 AAAGCTATATGCAAGATGGCATCTTAAAGAGACTTTGAAGAAATTAAGAAATA 190
OY 121 GCCAGCTGGAGATGTTGGCTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 180
DB 191 GCCAGCTGGAGATGTTGGCTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 250
OY 181 AACAGATCTATAGAGCATTTGGGCACTTCTGTTGATATCTGTTGAGCCAGACTGAGTGC 240
DB 251 AACAGATCTATAGAGCATTTGGGCACTTCTGTTGATATCTGTTGAGCCAGACTGAGTGC 310
OY 241 TCCAGAACTTGAAGAAAGCCATTCATATTTATGTAACAAAACCTGCGACGAAGATGGCTG 300
DB 311 TCCAGAACTTGAAGAAAGCCATTCATATTTATGTAACAAAACCTGCGACGAAGATGGCTG 369
OY 301 GAGAAAGTTCACTGGAGCCAGTGAAGTACCCCACTGGGAGAGAGGAGAAATCAGCT 360
DB 370 GAGAAAGTTCACTGGAGCCAGTGAAGTACCCCACTGGGAGAGAGGAGAAATCAGCT 429
OY 361 GTGATGCTGAGGCCAAGATTCATTAAGATAGCCATCTCTGGGCTTTGGCAGAGATTGGG 420
DB 430 GTGATGCTGAGGCCAAGATTCATTAAGATAGCCATCTCTGGGCTTTGGCAGAGATTGGG 489
OY 421 ACTCTTCAAGAGGCAATTAAGAGAAAGTTCTGGTGGTGAACCTTTTGCATGAATCTGAG 480
DB 490 ACTCTTCAAGAGGCAATTAAGAGAAAGTTCTGGTGGTGAACCTTTTGCATGAATCTGAG 549

```



```
OY      481 AGAAGGGCCCTGAGAAGCAAGAGGGAAAGTTGTGTTTATTAACCAACTTATCATCACTA   539
          |||||
Db       550 AGAAGGGCCTCAGAAAGCAAGAGGGAAAGATTGTGTTTATTAACCAACTTACAACCTCGA   608
```

Search completed: December 22, 2003, 20:22:18  
Job time : 5117.35 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 12:51:18 / Search time 341.284 Seconds  
(without alignments)  
11200.072 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_99\_1514  
Perfect score: 1416  
Sequence: 1 ATGAATTCCTTATCTTCG.....AGAAATGCTGCTAGTCC 1416

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: N\_Geneseq\_19Jun03.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1416	100.0	1851	19	AAV82779
2	1416	100.0	1851	24	ABO92016
3	1412.8	99.8	1767	21	AA40493
4	1412.8	99.8	1778	18	AAV02296
5	1412.8	99.8	1863	21	AAZ98034
6	1412.8	99.8	1863	22	AAD11647
7	1412.8	99.8	1863	24	ABK69743
8	1412.8	99.8	1863	25	ACCS0817

9	1412.8	99.8	1863	25	AB271453
10	1412.8	99.8	1884	21	AA258313
11	1412.8	99.8	1923	21	AA298139
12	1412.8	99.8	2077	25	ACCS0603
13	1412.8	99.8	2077	25	AB271331
14	1408	99.4	1860	22	AAK94491
15	1355.6	95.7	1895	22	AAH9703
16	1296.8	91.6	1784	23	AAH35592
17	1107.2	78.2	1778	24	ABK62598
18	862.8	59.5	895	21	AAA44359
19	772.4	54.5	1134	21	AAZ98101
20	772.4	54.5	1134	22	AAD11714
21	772.4	54.5	1134	24	ABK69810
22	772.4	54.5	1134	25	ACCS0818
23	772.4	54.5	1134	25	AB271454
24	520.6	36.8	609	20	AAV87487
25	470.2	33.2	642	22	AAK92181
26	470.2	33.2	642	22	AAK93447
27	284.8	20.1	314	20	AAV86369
28	276.6	19.5	462	18	AAV97398
29	276.6	19.5	462	18	AAV88060
30	276.6	19.5	462	18	AAV02139
31	242.4	17.1	441	22	AAK56831
32	212.6	15.0	317	20	AAK40587
33	114.2	8.1	492	22	AAK92932
34	90.8	6.4	394	14	AAO61260
35	84	5.9	217	20	AAK40583
36	61.4	4.3	424	23	AAK73591
37	60.2	4.3	65	24	ABN53183
38	60	4.2	60	24	ABN36212
39	42	3.0	784	22	AAH05496
40	42	3.0	2200	22	AAH15905
41	42	3.0	2200	22	AAH78055
42	42	3.0	2408	22	AAH99550
43	41	2.8	65	24	ABN28454
44	39.4	2.8	975	21	AACT7173
45	39.4	2.8	1556	24	AAJ33726

#### ALIGNMENTS

RESULT 1  
AAV82779  
ID AAV82779 standard; cDNA; 1851 BP.  
XX  
AC AAV82779;  
XX  
DT 25-FEB-1999 (first entry)  
XX  
DE Clone hu45\_2 isolated from human adult placenta cDNA library.  
XX  
KW Secreted protein; nutritional activity; immune stimulating; vaccine;  
KW Suppressing activity; haematopoiesis regulating activity;  
KW Tissue growth activity; activity; inhibin activity; chemotaxis;  
KW Chemokinetic activity; haemostasis; thrombolytic activity; receptor;  
KW Ligand; anti-inflammatory; cadherin; tumour invasion suppressor;  
KW Tumour inhibition; gene therapy; ds.  
XX  
OS Homo sapiens.  
XX  
PN M09842739-A2.  
XX  
PD 01-OCT-1998.  
XX  
PF 20-MAR-1998; 98MO-US05653.  
XX  
PR 19-MAR-1998; 98US-004446.  
XX  
PR 21-MAR-1997; 97US-0822167.  
XX  
PA (GENY) GENETICS INST INC.  
XX  
PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

PI Racie LA, Spaulding V, Treacy M;  
 XX WPI: 1998-609890/51.  
 DR P-PSDB; AAM83456.  
 XX New polynucleotides encoding secreted human proteins - derived from  
 PT human foetal brain, adult brain, foetal kidney, placenta or adult  
 PT pineal gland cDNA libraries.  
 XX  
 PS Claim 14; Page 69-70; 113pp; English.  
 CC The present sequence encodes a secreted protein. The polynucleotide and  
 CC secreted protein are predicted to have biological activities which would  
 CC make them suitable for treating, preventing or ameliorating medical  
 CC conditions in humans and animals, although no supporting data is given.  
 CC Suggested activities include nutritional activity, immune stimulating  
 CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating  
 CC activity, tissue growth activity, activin/inhibin activity,  
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,  
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour  
 CC invasion suppressor activity, and tumour inhibition activity (no data is  
 CC given in the specification to support these activities). The  
 CC polynucleotide is also stated to be useful for gene therapy.  
 XX  
 SQ Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;  
 Query Match 100.0%; Score 1416; DB 19; Length 1851;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGAATTCCTATCTTGGATTTTGGTGTTCACCTTTATCCCTGCTGCTGG 60  
 DB 99 ATGAATTCCTATCTTGGATTTTGGTGTTCACCTTTATCCCTGCTGCTGG 158  
 QY 61 AAAGTATATGCAAGATGGCATCTCTAAGAGACTTTTGAAGAAATGAAGAAATA 120  
 DB 159 AAAGTATATGCAAGATGGCATCTCTAAGAGACTTTTGAAGAAATGAAGAAATA 218  
 QY 121 GCCAGCTGTGAGATGTTGCTTAAGCATCATCACTAGCTGTTTATGTTAAAGCCAG 180  
 DB 219 GCCAGCTGTGAGATGTTGCTTAAGCATCATCACTAGCTGTTTATGTTAAAGCCAG 278  
 QY 181 AACAGATCTTATGAGCGATTTGGCACTTGTGTATCTGTGAGCCAGCTGAGTGGC 240  
 DB 275 AACAGATCTTATGAGCGATTTGGCACTTGTGTATCTGTGAGCCAGCTGAGTGGC 338  
 QY 241 TCCAAGAACTTGAAGAAAGCCATCCAAATTAATGTAACAAACCTGACCAAGATGGCTG 300  
 DB 339 TCCAAGAACTTGAAGAAAGCCATCCAAATTAATGTAACAAACCTGACCAAGATGGCTG 398  
 QY 301 GAGAAAGTTCACTGAGACCCAGTGAATTAACCCCACTGGAGAGGGGAGAAAGAACTCACT 360  
 DB 399 GAGAAAGTTCACTGAGACCCAGTGAATTAACCCCACTGGAGAGGGGAGAAAGAACTCACT 458  
 QY 361 GTGATGCTGAGCCAGAAATTCATTAAGATAGCCATCTGGGCTTGGGAGCGACTTGGG 420  
 DB 455 GTGATGCTGAGCCAGAAATTCATTAAGATAGCCATCTGGGCTTGGGAGCGACTTGGG 518  
 QY 421 ACTCTTCAGAGAGGATTAACAGAGAGTTCTGTGTGATCCTTTTGAATGAACCTGAG 480  
 DB 519 ACTCTTCAGAGAGGATTAACAGAGAGTTCTGTGTGATCCTTTTGAATGAACCTGAG 578  
 QY 481 AGAAGGGCTCAGAGAGGAGAAATGTTGTTTATTAACCACTTATCACTCACTAC 540  
 DB 579 AGAAGGGCTCAGAGAGGAGAAATGTTGTTTATTAACCACTTATCACTCACTAC 638  
 QY 541 TCAAGAGCGGTGCAATACGAAAGGAGGGGGGCTGAGAGCTGCCAAGGTGGGGCTTTG 600  
 DB 639 TCAAGAGCGGTGCAATACGAAAGGAGGGGGGCTGAGAGCTGCCAAGGTGGGGCTTTG 698  
 QY 601 GCATCTCTCATTCGATCCGCTGCTTCTTCATCTCACTCACTCACTCACTCACTCACT 660  
 DB 699 GCATCTCTCATTCGATCCGCTGCTTCTTCATCTCACTCACTCACTCACTCACTCACT 758

QY 661 GAATACAGAGATGGCGTCCCAAAATTCACAGCCCTGATTAAGGTGAGAGATGCAGAA 720  
 DB 759 GAATACAGAGATGGCGTCCCAAAATTCACAGCCCTGATTAAGGTGAGAGATGCAGAA 818  
 QY 721 ATGATGTCAGAAATGCGCTTCTCATGGGATCAAAATTTGCTCATCTTCAAGTAAAGTGGGCA 780  
 DB 819 ATGATGTCAGAAATGCGCTTCTCATGGGATCAAAATTTGCTCATCTTCAAGTAAAGTGGGCA 878  
 QY 781 AAGACTTACCCAGATTAATCTTCTTCAACACTGTACAGAGATCACTGGAGCAATAT 840  
 DB 879 AAGACTTACCCAGATTAATCTTCTTCAACACTGTACAGAGATCACTGGAGCAATAT 938  
 QY 841 CCAAGACAGGTTGATCTGTCAGTGAACATCTGAGACGCTGGAGTGTGGGACAGGCTGCC 900  
 DB 939 CCAAGACAGGTTGATCTGTCAGTGAACATCTGAGACGCTGGAGTGTGGGACAGGCTGCC 998  
 QY 901 ATGATGATGCGCGGTGAGGCTTTATATCATAGGAGCACTTCACTTATTAAGATCTT 960  
 DB 999 ATGATGATGCGCGGTGAGGCTTTATATCATAGGAGCACTTCACTTATTAAGATCTT 1058  
 QY 961 GGGCTGCTCCAAAGAGAGACTTGGCGCTGCTCTGAGCTGAGTGAAGAAAGAGTGA 1020  
 DB 1059 GGGCTGCTCCAAAGAGAGACTTGGCGCTGCTCTGAGCTGAGTGAAGAAAGAGTGA 1118  
 QY 1021 GTTGTCCTTCCAGTATTAATGTTACAGAGTAATATTTTCAACTCACTCACTGCTGCTG 1080  
 DB 1119 GTTGTCCTTCCAGTATTAATGTTACAGAGTAATATTTTCAACTCACTCACTGCTGCTG 1178  
 QY 1081 ATGAGATCTGACGAGAGAACCTTCTTACCACTGGGCTGCAATTCATGCGAGTGAAG 1140  
 DB 1179 ATGAGATCTGACGAGAGAACCTTCTTACCACTGGGCTGCAATTCATGCGAGTGAAG 1238  
 QY 1141 GCCAGGCGCATCATGAGAGAGTTATAGCTGCTGAGAGCCCTCATATATCATCTCACTG 1200  
 DB 1239 GCCAGGCGCATCATGAGAGAGTTATAGCTGCTGAGAGCCCTCATATATCATCTCACTG 1298  
 QY 1201 CTGAGCCATGAGAGAGAGGACAGACATCACTTTTGAATCCAGGTGAGTGGCTGAGAGCC 1260  
 DB 1299 CTGAGCCATGAGAGAGGACAGACATCACTTTTGAATCCAGGTGAGTGGCTGAGAGCC 1358  
 QY 1261 AGTCTACTGATGATTAATCAAGTATTTCTTCCATCATCTCCACGAGACACCATG 1320  
 DB 1359 AGTCTACTGATGATTAATCAAGTATTTCTTCCATCATCTCCACGAGACACCATG 1418  
 QY 1321 ACTGTATGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGCTGTTGTTCTTAT 1380  
 DB 1419 ACTGTATGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGCTGTTGTTCTTAT 1478  
 QY 1381 GTTGTCAGACATGAGAAAGATGCTGCTAGGCTC 1416  
 DB 1479 GTTGTCAGACATGAGAAAGATGCTGCTAGGCTC 1514

RESULT 2  
 ID AB092016 standard; cDNA; 1851 BP.  
 XX AB092016;  
 DT 04-OCT-2002 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 13.  
 XX  
 XX Human; cytosolic; antirheumatic; antiarthritic; vulnery; analgesic;  
 KW antinflammatory; antibacterial; immunosuppressive; antiparkinsonian;  
 KW neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;  
 KW antidiabetic; antidiabetic; antidiabetic; antidiabetic; antidiabetic;  
 KW immunostimulant; antiparasitic; secreted protein; transmembrane protein;  
 KW cytokine; cell proliferation; cell differentiation; autoimmune disease;  
 KW stem cell; growth factor; nervous system disease; neuropathy;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;  
 KW osteoporosis; severe combined immunodeficiency; SCID; infection;

KM multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.  
 XX Homo sapiens.  
 OS  
 XX US2002065394-A1.  
 PN  
 XX 30-MAY-2002.  
 PD  
 XX 22-DEC-2000; 2000US-0745763.  
 PF  
 XX 18-MAR-1998; 98US-0040963.  
 PR  
 XX (JACO/) JACOBS K.  
 PA (MCCO/) MCCOY J M.  
 PA (LAVA/) LAVALLIE E R.  
 PA (COLL/) COLLINS-RACIE L A.  
 PA (EVAN/) EVANS C.  
 PA (MERB/) MERBERG D.  
 PA (TREA/) TREACY M.  
 PA (SPAU/) SPAULDING V.  
 XX  
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Metberg D, Treacy M, Spaulding V;  
 XX  
 DR WPI; 2002-582343/62.  
 DR P-PSDB; ABP61800.  
 XX  
 PT Novel secreted or transmembrane protein and polynucleotide encoding the  
 PT protein, useful for diagnosis and treatment of neurological disorders,  
 PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis  
 PT  
 XX  
 PS Claim 50; Page 113-114; 284pp; English.  
 XX  
 CC The invention relates to human secreted or transmembrane protein (I),  
 CC their fragments and is encoded by specific complementary deoxyribonucleic  
 CC acid (cDNA) inserts (II), where the protein is substantially free from  
 CC other mammalian proteins. (I) are useful for preventing, treating or  
 CC ameliorating a medical condition, especially immunological treatment or  
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,  
 CC cytokine, cell proliferation, cell differentiation, anti-inflammatory,  
 CC stem cell growth factor activity and activin or inhibin-related  
 CC activities. (I) can be used to manipulate stem cells in culture to give  
 CC rise to neuroepithelial cells that can be used to augment or replace  
 CC cells damaged by illness, autoimmune disease, accidental damage or  
 CC genetic disorders. (I) induces the proliferation of neural cells and  
 CC regeneration of nerve and brain tissue and is useful for the treatment of  
 CC central and peripheral nervous system diseases and neuropathies, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic  
 CC activity, regulation of haematopoiesis and is useful for treating myeloid  
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopaenia  
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,  
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or  
 CC periodontal disease. (I) is also useful for gene protection or  
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury  
 CC in various tissues, various immune deficiencies and disorders including  
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,  
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,  
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,  
 CC such as asthma or other respiratory problems. (II) is useful to express  
 CC recombinant protein, as markers for tissues in which the corresponding  
 CC protein is preferentially expressed and in gene therapy. The present  
 CC sequence is that of a polynucleotide of the invention.  
 CC  
 XX  
 SQ Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;  
 Query Match 100.0%; Score 1416; DB 24; Length 1851;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 ATGAATTCCTTATCTTGCAATTTTCGGTGTGTTCACTTTATCCCTGTGCTGGG 60

Db  
 99 ATGAATTCCTTATCTTGCAATTTTCGGTGTGTTCACTTTATCCCTGTGCTGGG 158  
 Qy  
 61 AAGCTATATGCAAGATGCGATCTCTAAGAGACTTTTGAAGAAATTAAGAAATA 120  
 Db  
 159 AAGACTATATGCAAGATGCGATCTCTAAGAGACTTTTGAAGAAATTAAGAAATA 218  
 Qy  
 121 GCCAGCTGTGAGATGTTGCTTAAAGCAATCATCACTAGCTGTTTATGGTAAAGCCAG 180  
 Db  
 219 GCCAGCTGTGAGATGTTGCTTAAAGCAATCATCACTAGCTGTTTATGGTAAAGCCAG 278  
 Qy  
 181 AACAGATCTATAGCGATTTGGCACTTCTGTTTGAATCTGTTGACCCAGACTGAGTGGC 240  
 Db  
 279 AACAGATCTATAGCGATTTGGCACTTCTGTTTGAATCTGTTGACCCAGACTGAGTGGC 338  
 Qy  
 241 TCCAGAACTTAAGAAAAGCAATTCCTAATTTATGACCAAACTGCGACGAATGGGCTG 300  
 Db  
 339 TCCAGAACTTAAGAAAAGCAATTCCTAATTTATGACCAAACTGCGACGAATGGGCTG 398  
 Qy  
 301 GAGAAAGTTCACTGAGCGCAGTGAATATCCCACTGGGAGGAGGAGAAATCAGCT 360  
 Db  
 399 GAGAAAGTTCACTGAGCGCAGTGAATATCCCACTGGGAGGAGGAGAAATCAGCT 458  
 Qy  
 361 GTGATGCTGAGCGCAGTGAATATCATTAGATAGCCATCTGGGCTTTGGCAGCAGATTGGG 420  
 Db  
 459 GTGATGCTGAGCGCAGTGAATATCATTAGATAGCCATCTGGGCTTTGGCAGCAGATTGGG 518  
 Qy  
 421 ACTCTTCAAGAGCAATTAAGCAAGAAATTTCTGTTGATCTTTTCATGAACTGAC 480  
 Db  
 519 ACTCTTCAAGAGCAATTAAGCAAGAAATTTCTGTTGATCTTTTCATGAACTGAC 578  
 Qy  
 481 AGAAGGCTCTGAGAGCAAGAGGAGAAATTTGTTTATACCAACTTTCATCACTAC 540  
 Db  
 579 AGAAGGCTCTGAGAGCAAGAGGAGAAATTTGTTTATACCAACTTTCATCACTAC 638  
 Qy  
 541 TCAAGACGCTGCAATACCAAGCAGAGGGGCGGTGAGCTGCAAGTGGGGCTTTG 600  
 Db  
 639 TCAAGACGCTGCAATACCAAGCAGAGGGGCGGTGAGCTGCAAGTGGGGCTTTG 698  
 Qy  
 601 GCATCTCTCATGATCGTGGTCTCTTCCATCTCAAGTCTTCAACAGATATTCAG 660  
 Db  
 699 GCATCTCTCATGATCGTGGTCTCTTCCATCTCAAGTCTTCAACAGATATTCAG 758  
 Qy  
 661 GAATACAGAGTGGCGGCCAAATTCCAAGCCTGTATTCAGTGGGAAGATGCAAA 720  
 Db  
 759 GAATACAGAGTGGCGGCCAAATTCCAAGCCTGTATTCAGTGGGAAGATGCAAA 818  
 Qy  
 721 ATGATGTCAGAAATGAGCTTCTCATGAGATCAAAATTTGTCATTCAGTAAAGTGGGGCA 780  
 Db  
 819 ATGATGTCAGAAATGAGCTTCTCATGAGATCAAAATTTGTCATTCAGTAAAGTGGGGCA 878  
 Qy  
 781 AAGACTTACCAAGATGATGTTCTTCTCAACACTGTAGCAGAGATCACTGGGCAAAATAT 840  
 Db  
 879 AAGACTTACCAAGATGATGTTCTTCTCAACACTGTAGCAGAGATCACTGGGCAAAATAT 938  
 Qy  
 841 CCAGAAAGGTTGATGTCAGTGAAGCAATCTGCAAGCTGGGATGTTGGCAGAGGGCC 900  
 Db  
 939 CCAGAAAGGTTGATGTCAGTGAAGCAATCTGCAAGCTGGGATGTTGGCAGAGGGCC 998  
 Qy  
 901 ATGATGATGCGCGTGGAGCTTTATATCATGGAAGACACTCTCACTTATTAAGATCTT 960  
 Db  
 999 ATGATGATGCGCGTGGAGCTTTATATCATGGAAGACACTCTCACTTATTAAGATCTT 1058  
 Qy  
 961 GGGCTGCTCCAAAGAGAGACTCTGGGCTGGTCTGCACTGCAAGAAACAAGTGGGA 1020  
 Db  
 1059 GGGCTGCTCCAAAGAGAGACTCTGGGCTGGTCTGCACTGCAAGAAACAAGTGGGA 1118  
 Qy  
 1021 GTTGTGCTCTTCAAGTATTAATGATTACAGAGTAATATTTCCACTCAAGCTGGTGG 1080  
 Db  
 1119 GTTGTGCTCTTCAAGTATTAATGATTACAGAGTAATATTTCCACTCAAGCTGGTGG 1178  
 Qy  
 1081 ATGAGTCTGACGAGGAACTTTCTTACCACTGGGCTGCAATTCATGCGCAGTGAAG 1140

Db 1179 ATGAGATCTGACGACGAAACCTTTTACCCTAGGCTGCAATTCATGCGAGTGAAG 1238  
Qy 1141 GCCAGGGCCATCATGAGAGGTTATGAGCTGTGCGAGCCCTTCATATGACTCAGGTC 1200  
Db 1239 GCCAGGGCCATCATGAGAGGTTATGAGCTGTGCGAGCCCTTCATATGACTCAGGTC 1238  
Qy 1201 CTGAGCCATGAGAGAGGACAGACATCACTTTTGGATCAAGCTGAGTCCCTGAGACC 1260  
Db 1299 CTGAGCCATGAGAGAGGACAGACATCACTTTTGGATCAAGCTGAGTCCCTGAGACC 1358  
Qy 1261 AGTCTACTGATGATTAATACAGATTTTCTTCCATCACTCCACGAGACACCATG 1320  
Db 1359 AGTCTACTGATGATTAATACAGATTTTCTTCCATCACTCCACGAGACACCATG 1418  
Qy 1321 ACTGTCATGATCCAAAGACAGATGATGTGCTGCTGTTGGGCTGTTCTTAT 1380  
Db 1415 ACTGTCATGATCCAAAGACAGATGATGTGCTGCTGTTGGGCTGTTCTTAT 1478  
Qy 1381 GTTGTTCAGACATGAGAAATGCTGCTAGTCC 1416  
Db 1475 GTTGTTCAGACATGAGAAATGCTGCTAGTCC 1514

## RESULT 3

AAAA0493 standard; cDNA; 1767 BP.

AC AAA0493;

DT 16-NOV-2000 (first entry)

XX Human fetal kidney cDNA fragment AM282\_11.

XX Secreted protein; cytosolic; immunostimulatory; antimicrobial;  
XX antiviral; immunosuppressive; antiinflammatory; vulnery; cytokine;  
XX cell proliferation; differentiation; regulation; treatment; tumor;  
XX autoimmune disease; inflammatory disorder; wound; microbial infection;  
XX viral disease; graft versus host reaction suppression; ss.

OS Homo sapiens.

PN WO200037630-A1.

PD 29-JUN-2000.

XX 22-DEC-1999; 99WO-US31005.

XX 23-DEC-1998; 98US-0220876.

XX (GENY) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

PI Metberg D, Treacy M, Bowman MR;

XX MPI; 2000-442661/38.

DR P-PSDB; AAB10229.

XX Secreted human proteins AS296-11 and AS34-11, useful for treating

XX tumors, autoimmune diseases, inflammatory disorders, wounds, microbial

XX infections and viral diseases -

XX Disclosure, Page 198; 293pp; English.

XX This invention describes novel secreted human proteins (I) which have  
XX cytosolic, immunostimulatory, antimicrobial, antiviral,  
XX immunosuppressive, antiinflammatory and vulnery activity and which act  
XX as cytokine, cell proliferation or differentiation regulators. (I)  
XX is useful for treating tumors, autoimmune diseases, inflammatory  
XX disorders, wounds, microbial infections and viral diseases. (I) is also  
XX useful for suppressing graft versus host reaction. AAA0490-A40580  
XX represent cDNA fragments that encode the secreted proteins  
XX AAB10226-A10288 described in the method of the invention.

SQ Sequence 1767 BP; 512 A; 384 C; 418 G; 453 T; 0 other;

Query Match 99.8%; Score 1412.8; DB 21; Length 1767;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1414; Conservative 0; Indels 2; Gaps 0;

Qy 1 ATGAATTCCTTATCTTGGATTTTGGTGTGTTGACCTTTTATCCCTGTGCTGGG 60  
Db 6 ATGAATTCCTTATCTTGGATTTTGGTGTGTTGACCTTTTATCCCTGTGCTGGG 65  
Qy 61 AAAGCTATATGCAAGATGCGATCTTAAAGAGACTTTTGAAGAAATTAAGAAATA 120  
Db 66 AAAGCTATATGCAAGATGCGATCTTAAAGAGACTTTTGAAGAAATTAAGAAATA 125  
Qy 121 GCCAGCTGTGAGATGTTGCTTAAAGCAATCACTAGCTGTTATGTTAAAGCCAG 180  
Db 126 GCCAGCTGTGAGATGTTGCTTAAAGCAATCACTAGCTGTTATGTTAAAGCCAG 185  
Qy 181 AACAGATCTATGAGCCATTTGCACTTCTGTTGATACTGTGAGCCAGACTGAGTGC 240  
Db 186 AACAGATCTATGAGCCATTTGCACTTCTGTTGATACTGTGAGCCAGACTGAGTGC 245  
Qy 241 TCCAAAGACTTAAAGAAAGCCATCCAAATTAATGACCAAACTGACGAAATGGGCTG 300  
Db 246 TCCAAAGACTTAAAGAAAGCCATCCAAATTAATGACCAAACTGACGAAATGGGCTG 305  
Qy 301 GAGAAAGTTACCTGAGCCAGTGAATATCCCACTGGGAGAGAGGAGAAATCAGCT 360  
Db 306 GAGAAAGTTACCTGAGCCAGTGAATATCCCACTGGGAGAGAGGAGAAATCAGCT 365  
Qy 361 GTGATGTGAGCCAGAAATTCATTAATGATGATGATGATGATGATGATGATGATG 420  
Db 366 GTGATGTGAGCCAGAAATTCATTAATGATGATGATGATGATGATGATGATGATG 425  
Qy 421 ACTCTTCAAGAGCAATTAACGAGAAAGTTCTGTGTGATCTTTCATGATGATGATG 480  
Db 426 ACTCTTCAAGAGCAATTAACGAGAAAGTTCTGTGTGATCTTTCATGATGATGATG 485  
Qy 481 AGAAGGGCTCAG 540  
Db 486 AGAAGGGCTCAG 545  
Qy 541 TCAAGAGCGTGAATCCAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600  
Db 546 TCAAGAGCGTGAATCCAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 605  
Qy 601 GCATCTCTCATTCGATCGTGGCTCTTCTTCATCTCACTCACTCACTCACTCACT 660  
Db 606 GCATCTCTCATTCGATCGTGGCTCTTCTTCATCTCACTCACTCACTCACTCACT 665  
Qy 661 GAATACAGAGATGAGCGGCCAAATTCACAGCCTGATTAAGAGAGAGAGAGAGAG 720  
Db 666 GAATACAGAGATGAGCGGCCAAATTCACAGCCTGATTAAGAGAGAGAGAGAGAG 725  
Qy 721 ATGATGTCAAGATGAGCTTCTCATGAGATCAAAATTCATTCAGCTTAAAGAGAGAG 780  
Db 726 ATGATGTCAAGATGAGCTTCTCATGAGATCAAAATTCATTCAGCTTAAAGAGAGAG 785  
Qy 781 AAGACCTACCAAGATGAGATCTTCTTCAACACTGTAGAGAGATCACTGGAGAGAAAT 840  
Db 786 AAGACCTACCAAGATGAGATCTTCTTCAACACTGTAGAGAGATCACTGGAGAGAAAT 845  
Qy 841 CCAAGACAGGTTGATGAGTCAAGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 900  
Db 846 CCAAGACAGGTTGATGAGTCAAGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 905  
Qy 901 ATGATGATGAGCGGTGAGCTTATATCATGAGAGAGACTCTCACTTATTAAGATCTT 960  
Db 906 ATGATGATGAGCGGTGAGCTTATATCATGAGAGAGACTCTCACTTATTAAGATCTT 965  
Qy 961 GGGCTGCTCCAAAGAGAGACTCTGCGGTGCTCTGAGATGCAAGAAAGAGAGAGAG 1020  
Db 966 GGGCTGCTCCAAAGAGAGACTCTGCGGTGCTCTGAGATGCAAGAAAGAGAGAGAG 1025

QY 1021 GTTGGTCCCTTCAGTATTTATCAGTTACACAGGTAATATTTTCCAACTACAGTCTGGTG 1080  
 DB 1026 GTTGGTCCCTTCAGTATTTATCAGTTACACAGGTAATATTTTCCAACTACAGTCTGGTG 1085  
 QY 1081 ATGAGTCTGACGACGAGAACCTTTCTTACCCACTGCGCTGCAGTAATCTACGCGAGAAAG 1140  
 DB 1086 ATGAGTCTGACGACGAGAACCTTTCTTACCCACTGCGCTGCAGTAATCTACGCGAGAAAG 1145  
 QY 1141 GCCAGGCGCATCAGAGAGAGAGTATAGCCTGCTGCAGCCCCCTCAATATATCACTCAGTTC 1200  
 DB 1146 GCCAGGCGCATCAGAGAGAGAGTATAGCCTGCTGCAGCCCCCTCAATATATCACTCAGTTC 1205  
 QY 1201 CTGAGCCATGAGAGAGAGAGACATCACTTTTGGATCCAAAGCTGAGTCCGAGACC 1260  
 DB 1206 CTGAGCCATGAGAGAGAGAGACATCACTTTTGGATCCAAAGCTGAGTCCGAGACC 1265  
 QY 1261 AGTCTACTTGATGACTTATACAGATATTTCTTTCATCACTCCACGAGACACCATG 1320  
 DB 1266 AGTCTACTTGATGACTTATACAGATATTTCTTTCATCACTCCACGAGACACCATG 1325  
 QY 1321 ACTGTCATGATCCAAAGACATGAATGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1380  
 DB 1326 ACTGTCATGATCCAAAGACATGAATGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1385  
 QY 1381 GTTGTGACAGATGAGAAAGATGCTGCTAGTCC 1416  
 DB 1386 GTTGTGACAGATGAGAAAGATGCTGCTAGTCC 1421

RESULT 4  
 ID AAV02296 standard; DNA; 1778 BP.  
 AAV02296;

XX 21-MAY-1998 (first entry)  
 XX Human secreted protein AM282 full-length cDNA clone.  
 XX Secreted protein; AM282; cytokine; human; ds.  
 XX Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 17..1435  
 FT /\*tag= a  
 FT sig\_peptide 17..88  
 FT /\*tag= b  
 FT mat\_peptide 89..1432  
 FT /\*tag= b  
 XX  
 XX MO9739030-A2.  
 XX 23-OCT-1997.  
 XX  
 XX 16-APR-1997; 97WO-US06475.  
 XX  
 XX 13-JAN-1997; 97US-0783520.  
 XX 18-APR-1996; 96US-0634325.  
 XX  
 XX (GEMV ) GENETICS INST INC.  
 XX  
 XX Jacobs K, Lavallie ER, McCoy JM, Werberg D, Racie LA;  
 XX Spaulding V;  
 XX MPI; 1997-526400/48.  
 XX P-PSDB; AAM33604.  
 XX  
 XX New isolated secretory proteins AM340, AM282 and AK583 - possibly  
 XX have cytokine, cell proliferation/differentiation regulating,  
 XX immunomodulating activities, etc.

PS Claim 15; Page 44-45; 59pp; English.

XX This cDNA clone encodes a protein (see W33604) designated AM282.  
 CC It was identified as "y95b10.r1 human EST 30142.5" (GenBank  
 CC accession No. R77830) in a database search using a partial AM282  
 CC clone (see T97398) obtained from a human foetal kidney cDNA  
 CC library using methods selective for cDNAs encoding secreted  
 CC proteins. AM282 is deposited in ATCC 98026 together with clones  
 CC AM340 (see T97397) and AK583 (see V02297), which are also claimed.  
 CC AM282 protein can be used in a claimed method for preventing,  
 CC treating or ameliorating a medical condition. It may exhibit  
 CC cytokine, cell proliferation (either inducing or inhibiting) or  
 CC cell differentiation (either inducing or inhibiting) activity or  
 CC may induce production of other cytokines in certain cell  
 CC populations. It may also exhibit e.g. immune stimulating or  
 CC suppressing activity, haematopoiesis regulating activity, tissue  
 CC growth activity, activin/inhibin activity, chemokine or  
 CC chemokine activity, haemostatic or thrombolytic activity,  
 CC receptor/ligand activity, anti-inflammatory activity, tumour  
 CC inhibition activity, or other activities. No evidence of any of  
 CC these activities is given in the specification.

XX Sequence 1778 BP; 514 A; 386 C; 422 G; 456 T; 0 other;

Query Match 99.8%; Score 1412.8; DB 18; Length 1778;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAATTCCTTATCTTGGCATTTTGGTGGTGTCACTTTATCTCTGCTGGG 60  
 DB 17 ATGAATTCCTTATCTTGGCATTTTGGTGGTGTCACTTTATCTCTGCTGGG 76  
 QY 61 AAGCTATATGCAAGATGCGATCTCTAAGAGACCTTTGAAGAAATGAAGATA 120  
 DB 77 AAGCTATATGCAAGATGCGATCTCTAAGAGACCTTTGAAGAAATGAAGATA 136  
 QY 121 GCCAGCTGTGAGATGTTGCTAAGACATCACTAGCTGTTATGTTAAGCCAG 180  
 DB 137 GCCAGCTGTGAGATGTTGCTAAGACATCACTAGCTGTTATGTTAAGCCAG 196  
 QY 181 AACGATCTATGAGCATTTGGCATCTCTGTTGATACCTGTTGACCCAGCTAGTGC 240  
 DB 197 AACGATCTATGAGCATTTGGCATCTCTGTTGATACCTGTTGACCCAGCTAGTGC 256  
 QY 241 TCCAGAACCTAAGAAAGACATCCCAATATGATCCAAACCTGCGAGAAATGAGCTG 300  
 DB 257 TCCAGAACCTAAGAAAGACATCCCAATATGATCCAAACCTGCGAGAAATGAGCTG 316  
 QY 301 GAGAAAGTTCACTGAGCCAGTGAATACCCCACTGGGAGAGAGGAGAAATGAGCT 360  
 DB 317 GAGAAAGTTCACTGAGCCAGTGAATACCCCACTGGGAGAGAGGAGAAATGAGCT 376  
 QY 361 GTGATGCTGAGCCAGAAATTCATAGATGACATCTCTGCTGCTGAGAGATGAG 420  
 DB 377 GTGATGCTGAGCCAGAAATTCATAGATGACATCTCTGCTGCTGAGAGATGAG 436  
 QY 421 ACTCTCCAGAAAGCATTAAGAGAGATTTGTTGATGATCTTTGATGATGATGATG 480  
 DB 437 ACTCTCCAGAAAGCATTAAGAGAGATTTGTTGATGATCTTTGATGATGATGATG 496  
 QY 481 AGAAGGCTCAGAAAGAGAGAGATTTGTTGATGATCTTTGATGATGATGATGATG 540  
 DB 497 AGAAGGCTCAGAAAGAGAGAGATTTGTTGATGATCTTTGATGATGATGATGATG 556  
 QY 541 TCAAGACGCTGCAATACCGAAGCGAGGCGGTGAGAGCTGCAAGTGGGGCTTTTG 600  
 DB 557 TCAAGACGCTGCAATACCGAAGCGAGGCGGTGAGAGCTGCAAGTGGGGCTTTTG 616  
 QY 601 GCATCTCTCATTCATTCCTGCTCTCTTCTCCATCTACGCTCTCCACAGATGATGAG 660  
 DB 617 GCATCTCTCATTCATTCCTGCTCTCTTCTCCATCTACGCTCTCCACAGATGATGAG 676  
 QY 661 GAATACAGATGAGCTGCCAAATTCACAGCTGTATTAAGTGGAGATGACAGAA 720

```

Db 677 GAATACCAAGATGGGCTCCCAAGATTCACAGCTGTATTAGGTGGAGATGACAGAA 736
Qy 721 ATGATGCAAGATGGCTTCTCATGGGATCAAAATTTGCTACCTTAAGATGGGGCA 780
Db 737 ATGATGTCAGAAATGGCTTCTCATGGGATCAAAATTTGCTACCTTAAGATGGGGCA 796
Qy 781 AAGACCTACCAAGATGATGATCTTCTCAACCTGTAGAGAGATGATGATGGGAGCAATAT 840
Db 797 AAGACCTACCAAGATGATGATCTTCTCAACCTGTAGAGAGATGATGATGGGAGCAATAT 856
Qy 841 CCAGAACAGTTGTAATCTGTGATGAGATCTGACAGCTGGGATGTTGGGAGGGTCC 900
Db 857 CCAGAACAGTTGTAATCTGTGATGAGATCTGACAGCTGGGATGTTGGGAGGGTCC 916
Qy 901 ATGATGATGGCGGTGGAGCCCTTATATCATGGGAGCACTCTCATTTAAGATCTT 960
Db 917 ATGATGATGGCGGTGGAGCCCTTATATCATGGGAGCACTCTCATTTAAGATCTT 976
Qy 961 GGGGCTGCTCCAAAGAGACTCTGGGCTGGTCTGGAATGCAAGAAACAAGTGA 1020
Db 977 GGGGCTGCTCCAAAGAGACTCTGGGCTGGTCTGGAATGCAAGAAACAAGTGA 1036
Qy 1021 GTTGGTGCCTTCCAGTATTTATCATGTAACAAGTAAATTTCCACTACAGTCTGTTG 1080
Db 1037 GTTGGTGCCTTCCAGTATTTATCATGTAACAAGTAAATTTCCACTACAGTCTGTTG 1096
Qy 1081 ATGAGTCTGACGCAAGAACTTCTTACCCACTGGGCTGCAATTACTGGCAGTGAAG 1140
Db 1097 ATGAGTCTGACGCAAGAACTTCTTACCCACTGGGCTGCAATTACTGGCAGTGAAG 1156
Qy 1141 GCCAGGGCCATCATGAGAGAGTTATGAGCCCTGCTGAGCCCTCAATATCATCTAGGTC 1200
Db 1157 GCCAGGGCCATCATGAGAGAGTTATGAGCCCTGCTGAGCCCTCAATATCATCTAGGTC 1216
Qy 1201 CTGAGCCATGAGAGAGAGCAACATCACTTTTGGATCAAGCTGAGTCTGAGACC 1260
Db 1217 CTGAGCCATGAGAGAGAGCAACATCACTTTTGGATCAAGCTGAGTCTGAGACC 1276
Qy 1261 AGTCTACTGATGATCTTATCAAGATATTTCTTCTCATCTCCACGAGACACCATG 1320
Db 1277 AGTCTACTGATGATCTTATCAAGATATTTCTTCTCATCTCCACGAGACACCATG 1336
Qy 1321 ACTGTCATGATCCAAAGAGATGATGATGCTGCTGCTGTTGGCTGTTCTTAT 1380
Db 1337 ACTGTCATGATCCAAAGAGATGATGATGCTGCTGCTGTTGGCTGTTCTTAT 1396
Qy 1381 GTTGTGCAAGATGAGAAATGCTGCTAGTCC 1416
Db 1397 GTTGTGCAAGATGAGAAATGCTGCTAGTCC 1432

```

RESULT 5  
AAZ98034  
ID AAZ98034 standard; cDNA; 1863 BP.

AC AAZ98034;  
DT 09-MAY-2000 (first entry)

DE Human secreted protein encoding nucleotide sequence SEQ ID NO:28.

XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;  
XX antiinflammatory; neurotropic; neuroprotective; antiallergic; cancer;  
XX tumour; neurodegenerative disorder; developmental abnormality; allergy;  
XX foetal deficiency; blood disorder; immune system disorder; arthritis;  
XX autoimmune disease; hepatic disease; renal disease; inflammation;  
XX Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;  
XX infection; AIDS; spinal cord injury; transplant rejection; diabetes;  
XX asthma; sepsis; acne; psoriasis; cardiovascular disorder;  
XX reproductive disorder; gastrointestinal disorder; respiratory disorder;  
XX metabolic disorder; food additive; preservative; ss.

```

OS Homo sapiens.
XX
XX WO200004140-A1.
XX
XX 27-JAN-2000.
XX
XX 14-JUL-1999; 99WO-US15849.
XX
XX 15-JUL-1998; 98US-0092921.
XX 15-JUL-1998; 98US-0092922.
XX 15-JUL-1998; 98US-0092956.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
XX Lalleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE;
XX Muenzel M, Endress GA, Soppet DR;
XX
XX WPI; 2000-161128/14.
XX
XX P-PSDB; AAY87081.
XX
XX New isolated human genes, useful for diagnosis and treatment of, e.g.
XX cancers, neurological or blood disorders
XX
XX Claim 1; Page 319; 494pp; English.
XX
XX The polynucleotide sequences given in AAZ98017 to AAZ98108 encode the
XX human secreted proteins given in AAY87064 to AAY87223. Human secreted
XX protein can have activities based on the tissues and cells the genes are
XX expressed in. Examples of activities include: cytostatic;
XX immunosuppressive; antiinflammatory; neurotropic; neuroprotective; and
XX antiallergic. The polynucleotides and their corresponding secreted
XX polypeptides are useful for preventing, treating or ameliorating medical
XX conditions, e.g. by protein or gene therapy. Also pathological conditions
XX can be diagnosed by determining the amount of the new polypeptides in a
XX sample or by determining the presence of mutations in the new
XX polynucleotides. Human secreted proteins and their polynucleotides can
XX be used for developing products for the diagnosis or treatment of cancer,
XX tumours, neurodegenerative disorders, developmental abnormalities and
XX foetal deficiencies, blood disorders, diseases of the immune system,
XX autoimmune diseases, hepatic and renal disease, inflammation,
XX allergies, Alzheimer's disease, behavioural disorder, schizophrenia,
XX osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
XX transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
XX cardiovascular disorders, reproductive disorders, gastrointestinal
XX disorders, respiratory disorders and metabolic disorders. The
XX proteins or polynucleotides can also be used as food additives or
XX preservatives. The proteins are also useful for identifying their
XX binding partners. AAZ98008 to AAZ98016 and AAY87063 are sequence used in
XX the exemplification of the present invention.
XX
XX Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

```

Query Match 99.8%; Score 1412.8; DB 21; Length 1863;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 ATGAAATTCCTTATCTTGGATTTTGGTGTGATCACTTTATCCCTGTGCTGGG 60
Db 99 ATGAAATTCCTTATCTTGGATTTTGGTGTGATCACTTTATCCCTGTGCTGGG 158
Qy 61 AAGCTATATGCAAGATGGCATCTTAAGAGACTTTTGAAGAAATTAAGAGAAATA 120
Db 159 AAGCTATATGCAAGATGGCATCTTAAGAGACTTTTGAAGAAATTAAGAGAAATA 218
Qy 121 GCCAGCTGTGAGATGTTGCTTAAGACATCACTAAGCTGTTATGTAAGCCGAG 180
Db 219 GCCAGCTGTGAGATGTTGCTTAAGACATCACTAAGCTGTTATGTAAGCCGAG 278
Qy 181 AACGATCTTATGAGCGATTTGGCTTGTGATCTGTTGAGCCGAGCTGAGTGGC 240
Db 279 AACGATCTTATGAGCGATTTGGCTTGTGATCTGTTGAGCCGAGCTGAGTGGC 338

```



QY 241 TCAGAGACCTAGAAAAAGCCATCCAAATTATGTAACAAAACTGACAGAGATGGGCTG 300  
 DB 339 TCAGAGACCTAGAAAAAGCCATCCAAATTATGTAACAAAACTGACAGAGATGGGCTG 398  
 QY 301 GAGAAAGTTCACTGAGAGCCAGTGAATACCCACTGGAGAGAGGAGAAATCAGCT 360  
 DB 399 GAGAAAGTTCACTGAGAGCCAGTGAATACCCACTGGAGAGAGGAGAAATCAGCT 458  
 QY 361 GTGATGCTGGAGCCAGAAATTCATTAAGTAGCAATCCGGGCTTGGGAGAGAGATGGG 420  
 DB 459 GTGATGCTGGAGCCAGAAATTCATTAAGTAGCAATCCGGGCTTGGGAGAGAGATGGG 518  
 QY 421 ACTCTCTCAGAAAGCATTACAGAGAAAGTTGATGATGATCTCTTTCATGATGATGATG 480  
 DB 519 ACTCTCTCAGAAAGCATTACAGAGAAAGTTGATGATGATCTCTTTCATGATGATGATG 578  
 QY 481 AGAAGGCTCTCAGAAAGCAGAGAGAGATTTGTTTATTAACCACTTATCATCACTAC 540  
 DB 579 AGAAGGCTCTCAGAAAGCAGAGAGAGATTTGTTTATTAACCACTTATCATCACTAC 638  
 QY 541 TCAGAGAGGTCATACCGAAGCGAGGGGGGGGAGGTCGCAAGGTTGGGGGCTTTG 600  
 DB 639 TCAGAGAGGTCATACCGAAGCGAGGGGGGGGAGGTCGCAAGGTTGGGGGCTTTG 698  
 QY 601 GCATCTCTCATTCGATCGTGGCTCTCTCTCTCATCTACAGTCTCTCAACAGATATTCAG 660  
 DB 699 GCATCTCTCATTCGATCGTGGCTCTCTCTCTCATCTACAGTCTCTCAACAGATATTCAG 758  
 QY 661 GAATACCAAGATGCGTCCCAAAATTCAAACAGCTGATTAACGTTGGAAGATGCAAG 720  
 DB 759 GAATACCAAGATGCGTCCCAAAATTCAAACAGCTGATTAACGTTGGAAGATGCAAG 818  
 QY 721 ATGATGTCAGAAAGTGGCTCTCTCATGAGATCAAAATTCATTCAGCTTAAAGATGGGGCA 780  
 DB 819 ATGATGTCAGAAAGTGGCTCTCTCATGAGATCAAAATTCATTCAGCTTAAAGATGGGGCA 878  
 QY 781 AAGACTTACCAAGATGATGATCTCTTCAACACTGAGAGAGATCACTGGAGCAAAATAT 840  
 DB 879 AAGACTTACCAAGATGATGATCTCTTCAACACTGAGAGAGATCACTGGAGCAAAATAT 938  
 QY 841 CCAGAACAGTTGTAATCTGTCAGTGAATTCGACAGCTGGAGTGTGGGCAAGGCTGCC 900  
 DB 939 CCAGAACAGTTGTAATCTGTCAGTGAATTCGACAGCTGGAGTGTGGGCAAGGCTGCC 998  
 QY 901 ATGATGATGAGCGGTGGAGCCTTATATATCATGAGAGAGACTCTCACTTATTAAGATCTT 960  
 DB 999 ATGATGATGAGCGGTGGAGCCTTATATATCATGAGAGAGACTCTCACTTATTAAGATCTT 1058  
 QY 961 GGGCTGCGTCCAAAGAGAGACTCTGGGCGTGGTCTCTGGACTGCAAGAGAACAGGTGGA 1020  
 DB 1059 GGGCTGCGTCCAAAGAGAGACTCTGGGCGTGGTCTCTGGACTGCAAGAGAACAGGTGGA 1118  
 QY 1021 GTTGGTGGCTTCCAGATATTAATCAAGTGAAGGTAATTTTCCAACTACAGTCTGGTG 1080  
 DB 1119 GTTGGTGGCTTCCAGATATTAATCAAGTGAAGGTAATTTTCCAACTACAGTCTGGTG 1178  
 QY 1081 ATGAGATGAGCGAGAGAACTCTTACCACTGGGCTGCAATTAATCACTGGAGAGAAAG 1140  
 DB 1179 ATGAGATGAGCGAGAGAACTCTTACCACTGGGCTGCAATTAATCACTGGAGAGAAAG 1238  
 QY 1141 GCCAGGGCCATCATGAGAGAGGTTATGAGCCTGCTGCGAGCCCTCAATATCATCAAGTTC 1200  
 DB 1239 GCCAGGGCCATCATGAGAGAGGTTATGAGCCTGCTGCGAGCCCTCAATATCATCAAGTTC 1298  
 QY 1201 CTGAGCCATGAGAGAGGAGACATCAACTTTTGGATCCAAGCTGAGAGTGGTGGAGCC 1260  
 DB 1299 CTGAGCCATGAGAGAGGAGACATCAACTTTTGGATCCAAGCTGAGAGTGGTGGAGCC 1358  
 QY 1261 AGTCTACTGATGATTAACAAGTAATTTCTTCTTCATCACTCCCAAGAGAGACCACTG 1320  
 DB 1359 AGTCTACTGATGATTAACAAGTAATTTCTTCTTCATCACTCCCAAGAGAGACCACTG 1418  
 QY 1321 ACTGTCATGATCCAAAGCAGATGATGCTGCTGTTGGGCTGTTGTTCTTAT 1380

DB 1419 ACTGTCATGATCCAAAGCAGATGAATGTTGCTGCTCTGTTGGCTGTTGTTCTTAT 1478  
 QY 1381 GTTCTGACAGATGAGAGAAATGCTGCTTGGTCC 1416  
 DB 1479 GTTCTGACAGATGAGAGAAATGCTGCTTGGTCC 1514  
 RESULT 6  
 AAD1647  
 ID AAD1647 standard; cDNA; 1863 BP.  
 XX  
 AC AAD1647;  
 DT 24-SEP-2001 (first entry)  
 XX  
 DE Human secreted protein-encoding gene 18 cDNA clone HRA3J35, SEQ ID NO:28.  
 KW Human; secreted protein; proliferative disorder; cancer; tumour; asthma;  
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 KW parkinson's disease; cognitive disorder; schizophrenia; skin disorder;  
 KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;  
 KW inflammation; neurological disorder; Alzheimer's disease; food additive;  
 KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;  
 KW pregnancy-related disorder; endocrine disorder; infection; wound healing;  
 KW cell culture; chemotaxis; vulnery; binding partner identification;  
 KW gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 99..1517 /tag= a  
 FT CDS /product= "Human secreted protein precursor"  
 FT s1g\_peptide 99..170 /tag= b  
 FT 171..1514 /tag= c  
 FT mat\_peptide /tag= c  
 FT /product= "Mature human secreted protein"  
 FT  
 PN MO200151504-A1.  
 PN  
 PD 19-JUL-2001.  
 PD  
 PF 12-JAN-2001; 2001WO-US00911.  
 PF  
 PR 13-JAN-2000; 2000US-0482273.  
 PR  
 PA (HOMA-) HUMAN GENOME SCI INC.  
 PA  
 PI Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y;  
 PI Lalleur DW, Olesen HS, Brewer LA, Florence KA, Young PE, Soppet DR;  
 PI Endress GA, Muscenski M, Ebner R;  
 XX  
 DR WPI: 2001-425865/45.  
 DR P-PSDB; AAB06058.  
 DR  
 PT Isolated nucleic acid molecule encoding a human secreted protein is  
 PT used in preventing, treating or ameliorating a medical condition -  
 PT  
 XX  
 PS Claim 1; Page 686; 864pp; English.  
 PS  
 CC AAD1630-AAD11721 represent cDNAs corresponding to 71 human secreted  
 CC protein genes, and AAB06041-AAB06132 represent the proteins they encode.  
 CC AAB06133-AAB06205 represent human secreted protein fragments.  
 CC The secreted proteins and their genes are useful for preventing, treating  
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
 CC Pathological conditions can be diagnosed, e.g., by determining the amount of the  
 CC new protein in a sample or by determining the presence of mutations in  
 CC the new genes. Specific uses are described for each of the 71 genes,  
 CC based on the tissues in which they are most highly expressed, and include  
 CC developing products for the diagnosis or treatment of proliferative



XX	04-APR-2002.
PD	
XX	
XX	24-SEP-2001; 2001WO-US29871.
XX	
PR	25-SEP-2000; 2000US-234925P.
PR	12-JAN-2001; 2001WO-US00911.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PI	Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y;
PI	Lafleur DM, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;
PI	Endress GA, Mucenski M, Ebner R;
DR	
XX	WPI; 2002-362489/39.
DR	P-PSDB; ABG33880.
XX	
PT	Novel 71 isolated secreted polypeptides and polynucleotides encoding
PT	the polypeptides, useful for treating Huntington's disease, sepsis,
PT	meningitis, thrombocytopenia, haemolytic anaemia, rheumatoid arthritis,
XX	aschma
PS	
XX	Claim 1; Page 1170; 1478pp; English.
CC	The invention relates to an isolated nucleic acid molecule (or its
CC	fragment, homologue complement or allelic variant) encoding a human
CC	secreted protein (and its fragment, domain, epitope, variant, secreted
CC	form and species variant). Also included are a recombinant vector
CC	comprising the nucleic acid, a recombinant host cell comprising the
CC	vector, an antibody against the secreted protein, a recombinant host cell
CC	that expresses the secreted protein and a method of identifying a binding
CC	partner of the secreted protein. The nucleic acid and protein are used to
CC	prevent, diagnose, treat or ameliorate a medical condition in e.g.
CC	humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep
CC	for example autoimmune diseases e.g. rheumatoid arthritis,
CC	hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC	cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC	e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC	Alzheimer's disease, infections caused by bacteria, viruses and fungi and
CC	ocular disorders e.g. corneal infection. Many other diseases and
CC	disorders are listed in the specification. The polypeptides can also be
CC	used to aid wound healing an epithelial cell proliferation, to prevent
CC	skin aging due to sunburn, to maintain organs before transplantation, for
CC	supporting cell culture of primary tissues, to regenerate tissues and in
CC	chemotherapy. The polypeptides can also be used as a food additive or
CC	preservative to increase or decrease storage capabilities. The present
CC	sequence encodes a novel human secreted protein of the invention.
XX	
SQ	Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;
Query Match	99.8%; Score 1412.8; DB 24; Length 1863;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 1414; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
OY	1 ATGAATTCCTTATCTCGCATTTTGGGTGTGTCACCTTTATCCCTGTCTGGG 60
Db	99 ATGAATTCCTTATCTCGCATTTTGGGTGTGTCACCTTTATCCCTGTCTGGG 158
OY	61 AAAGCTAATGCAGAATAATGCGACTCTTAAGAGACCTTTGAAAGAAATAAAGACAATA 120
Db	159 AAAGCTAATGCAGAATAATGCGACTCTTAAGAGACCTTTGAAAGAAATAAAGACAATA 218
OY	121 GCCAGCTGTGAGATGTTGCTTAAGCAATCATTACCTAGCTGTTTAGTAAAGCCGAC 180
Db	219 GCCAGCTGTGAGATGTTGCTTAAGCAATCATTACCTAGCTGTTTAGTAAAGCCGAC 278
OY	181 AACAGATCCTATGAGCGATTGGCACCTTGCTGTGATGTAAGTGTGAAGCCAGACTAGTGGC 240
Db	279 AACAGATCCTATGAGCGATTGGCACCTTGCTGTGATGTAAGTGTGAAGCCAGACTAGTGGC 338
OY	241 TCAGAAGACTTGAAGAAAAGCCATCCAATTAATGTCACAAAACCTGACGACAGATGGCGTG 300
Db	339 TCAGAAGACTTGAAGAAAAGCCATCCAATTAATGTCACAAAACCTGACGACAGATGGCGTG 398

QY	301	TAGAAAGTTCCCTCGAGCCAGTGAATAATCCCACTGGGAGAGGGAGAAAGATCACT	360
Db	399	GAGAAAGTTCACTCGAGCCAGTGAATAATCCCACTGGGAGAGGGAGAAAGATCACT	458
QY	361	GTGATGCTGGAGCCAAAGATTCAATAGATACCATCTGGGCTTTGGACGACATTGGG	420
Db	459	GTGATGCTGGAGCCAAAGATTCAATAGATACCATCTGGGCTTTGGACGACATTGGG	518
QY	421	ACTCCTCAGAAAGGCATTACAGCAGAAATTCTGGTGTGACCTCTTTCGATGAACCTGCAG	480
Db	519	ACTCCTCAGAAAGGCATTACAGCAGAAAGTTCTGGTGTGACCTCTTTCGATGAACCTGCAG	578
QY	481	AGAAAGGCTCTAGAAAGCAGAGGGAGATTGTGTTATTAACCAACTTACATCACTAC	540
Db	579	AGAAAGGCTCTAGAAAGCAGAGGGAGATTGTGTTATTAACCAACTTACATCACTAC	638
QY	541	TCAAGGACGGGTCAATACCGAAGCAGGGGGGGGTGGAAAGCTGCAGAGTGGGGGCTTTG	600
Db	639	TCAAGGACGGGTCAATACCGAAGCAGGGGGGGGTGGAAAGCTGCAGAGTGGGGGCTTTG	698
QY	601	GCATCTCTCATTCGATCCGTGGCCCTCTCTCCATCTACAGTCCCTCACACAGTATTGAG	660
Db	699	GCATCTCTCATTCGATCCGTGGCCCTCTCTCCATCTACAGTCCCTCACACAGTATTGAG	758
QY	661	GAATPACCAAGATGCGTGCCTCCAAATTTCCACACAGCTGTATTTACGGTGAAGATGCAGAA	720
Db	759	GAATPACCAAGATGCGTGCCTCCAAATTTCCACACAGCTGTATTTACGGTGAAGATGCAGAA	818
QY	721	ATGATATTCANAATAATGGCTTCTCATGGGATCAAAATTGTCACTTACGCTPAAACATGGGGCA	780
Db	819	ATGATATTCANAATAATGGCTTCTCATGGGATCAAAATTGTCACTTACGCTPAAACATGGGGCA	878
QY	781	AAGACTTACCAGATCTAGTATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATAT	840
Db	879	AAGACTTACCAGATCTAGTATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATAT	938
QY	841	CCAGAACAGGTTGTACTGTGCAGTGGACATCTGACACAGCTGGGATGTTTGGCAGGGTCC	900
Db	939	CCAGAACAGGTTGTACTGTGCAGTGGACATCTGACACAGCTGGGATGTTTGGCAGGGTCC	998
QY	901	ATGATATGATGGCGGTGGAGCTTTATATCATGGGAAGCATCTCACTTATTTAAAGATCTT	960
Db	999	ATGATATGATGGCGGTGGAGCTTTATATCATGGGAAGCATCTCACTTATTTAAAGATCTT	1058
QY	961	GGGCGCGCTCCAAAGAGGACTCTGCGGGTGGTCTGGACTGGAGAAAGCAAGGTGGA	1020
Db	1059	GGGCGCGCTCCAAAGAGGACTCTGCGGGTGGTCTGGACTGGAGAAAGCAAGGTGGA	1118
QY	1021	GTGGTGGCTTCCAGTATTTATCAGTTTACACAGGTAAATATTTCCACTACAGTCTGGTG	1080
Db	1119	GTGGTGGCTTCCAGTATTTATCAGTTTACACAGGTAAATATTTCCACTACAGTCTGGTG	1178
QY	1081	ATGAGATCTGACGAGGAACCTTCTTACCACTGCGCTGCAAATTCATCTGGCAGTGAAGA	1140
Db	1179	ATGAGATCTGACGAGGAACCTTCTTACCACTGCGCTGCAAATTCATCTGGCAGTGAAGA	1238
QY	1141	GCCAGGGCCATCAGAGAGAGAGTATAGACCTGCTGCAGCCCCCTCAATATCACTCAAGCT	1200
Db	1239	GCCAGGGCCATCAGAGAGAGAGTATAGACCTGCTGCAGCCCCCTCAATATCACTCAAGCT	1298
QY	1201	CTGAGCCATGAGAGAGGACAGACATCACTTTTGGATCCAAAGCTGAGTCCCTGGAGCC	1260
Db	1299	CTGAGCCATGAGAGAGGACAGACATCACTTTTGGATCCAAAGCTGAGTCCCTGGAGCC	1358
QY	1261	AGTCTACTTGAATCTTATACAGATATTTCTTCTTCATCACTCCACGAGACACCATG	1320
Db	1359	AGTCTACTTGAATCTTATACAGATATTTCTTCTTCATCACTCCACGAGACACCATG	1418
QY	1321	ACTGTCAATGATCCAAAGCAGATGAATGTGCTGTCTGTTGGGCTGTGTTCTTAT	1380
Db	1419	ACTGTCAATGATCCAAAGCAGATGAATGTGCTGTCTGTTGGGCTGTGTTCTTAT	1478

QY 1381 GTGTTCAGACATGAGAAATGCTGCTAGTCC 1416  
DB 1479 GTGTTCAGACATGAGAAATGCTGCTAGTCC 1514

RESULT 8  
ACC50817  
ID ACC50817 standard; cDNA; 1863 BP.  
XX  
AC ACC50817;  
XX  
DT 12-JUN-2003 (first entry)  
XX  
DE Human secreted protein coding sequence, SEQ ID 484.  
XX  
KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytosstatic;  
KW vulnerability; antiinflammatory; neurotropic; neuroprotective;  
KW antiparkinsonian; gene therapy; human; cardiovascular disorder;  
KW gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200295010-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 19-MAR-2002; 2002WO-US09785.  
XX  
PR 21-MAR-2001; 2001US-277340P.  
PR 19-JUL-2001; 2001US-306171P.  
PR 13-NOV-2001; 2001US-331287P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM;  
XX  
DR WPI; 2003-129429/12.  
XX  
PT Novel human secreted proteins, useful for detecting, preventing,  
PT diagnosing, prognosticating, treating and/or ameliorating  
PT cardiovascular disorders such as arrhythmia -  
XX  
XX  
XX Claim 21; SEQ ID 484; 1881bp; English.  
XX  
XX The present invention relates to novel human secreted proteins  
XX (ABR4733-ABR48145) and their coding sequences (ACC50744-ACC50856). The  
XX proteins and their coding sequences are useful for the preparation of a  
XX diagnostic or pharmaceutical composition for diagnosing or treating a  
XX cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,  
XX coronary arteriosclerosis and myocardial ischaemia), neural disorders,  
XX immune system disorders, muscular disorders, reproductive disorders,  
XX gastrointestinal disorders, pulmonary disorders, renal disorders,  
XX proliferative disorders and/or cancerous diseases and conditions, for  
XX wound healing and epithelial cell proliferation, to treat inflammation or  
XX infection, for treating thrombosis and arteriosclerosis, for treating or  
XX preventing neural damage which occurs in neuronal disorders or  
XX neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
XX disease, to enhance bone and periodontal regeneration and aid in tissue  
XX transplants or bone grafts, to prevent skin aging or hair loss, to  
XX stimulate growth and differentiation of haematopoietic cells and bone  
XX marrow cells when used in combination with other cytokines, to maintain  
XX organs before transplantation or for supporting cell culture of primary  
XX tissues, to increase or decrease differentiation or proliferation of  
XX embryonic stem cells, or to modulate mammalian characteristics or  
XX metabolism.  
XX Note: The sequence data for this patent was published in electronic  
XX format and is available from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Query March 99.8%; Score 1412.8; DB 25; Length 1863;  
Best Local Similarity 99.9%; Pred. No. 0;

Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAATTCCTTATCTGGATTTTGGTGGTTCACCTTTATCCCTGCTGGG 60  
DB 99 ATGAATTCCTTATCTGGATTTTGGTGGTTCACCTTTATCCCTGCTGGG 158  
QY 61 AAAGCTATATGCAAGAAATGGCATCTTAAGAGCACTTTTGAAGAAATGAAGAAATA 120  
DB 159 AAAGCTATATGCAAGAAATGGCATCTTAAGAGCACTTTTGAAGAAATGAAGAAATA 218  
QY 121 GCCAGCTGTGAGAGATGTTGCTTAAGCAATCATCACTAGCTGTTTATGTTAAGCCCA 180  
DB 219 GCCAGCTGTGAGAGATGTTGCTTAAGCAATCATCACTAGCTGTTTATGTTAAGCCCA 278  
QY 181 AACGATCTTATGAGCGATTTGGCACTCTGCTGTTATCTGTGAGCCAGCTAGTGGC 240  
DB 279 AACGATCTTATGAGCGATTTGGCACTCTGCTGTTATCTGTGAGCCAGCTAGTGGC 338  
QY 241 TCCAAAGAACTTGAAGAAAGCCATCCAAATTAATGTAACCAAACTGACAGAAATGGGCTG 300  
DB 339 TCCAAAGAACTTGAAGAAAGCCATCCAAATTAATGTAACCAAACTGACAGAAATGGGCTG 398  
QY 301 GAGAAAGTTCACTGAGAGCCAGTAAGTAATCCCACTGGAGAGGGAGAAATCACT 360  
DB 399 GAGAAAGTTCACTGAGAGCCAGTAAGTAATCCCACTGGAGAGGGAGAAATCACT 458  
QY 361 GTGATGCTGAGAGCCAAATTCATTAAGTACCATCTGCTGCTGGACAGCATGGG 420  
DB 459 GTGATGCTGAGAGCCAAATTCATTAAGTACCATCTGCTGCTGGACAGCATGGG 518  
QY 421 ACTCTCCAGAGGCACTTACAGCAGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB 519 ACTCTCCAGAGGCACTTACAGCAGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 578  
QY 481 AGAAGGCTCAGAGGCAAGGAGGAGATGTTGTTTAAACCACTTACATCACTAC 540  
DB 579 AGAAGGCTCAGAGGCAAGGAGGAGATGTTGTTTAAACCACTTACATCACTAC 638  
QY 541 TCAAGAGCGGTCATACCAAGCGAGGGGCGGTGAGGTGCAAGGTGGGGGCTTTG 600  
DB 639 TCAAGAGCGGTCATACCAAGCGAGGGGCGGTGAGGTGCAAGGTGGGGGCTTTG 698  
QY 601 GCATCTCTCATTCGATCGTGGCTCTCTCTCATCTACAGTCTCTCACAGGATTTAC 660  
DB 699 GCATCTCTCATTCGATCGTGGCTCTCTCTCATCTACAGTCTCTCACAGGATTTAC 758  
QY 661 GAATACAGAGATGGCGTGGCCCAAAATTCACACGCTGTTATTCGTTGAGATGCAGAA 720  
DB 759 GAATACAGAGATGGCGTGGCCCAAAATTCACACGCTGTTATTCGTTGAGATGCAGAA 818  
QY 721 ATGATGTCAGAAATGGCTTCTCATGAGATCAAAATGCTCATTCAGTTAAAGATGGGGCA 780  
DB 819 ATGATGTCAGAAATGGCTTCTCATGAGATCAAAATGCTCATTCAGTTAAAGATGGGGCA 878  
QY 781 AAGACTTACCAAGATCTGATTCCTTCAACACTGTACAGAGATCACTGGGAGCAATAT 840  
DB 879 AAGACTTACCAAGATCTGATTCCTTCAACACTGTACAGAGATCACTGGGAGCAATAT 938  
QY 841 CCAGAACAGGTTGATGCTGATGAGCATCTGACAGCTGGAGTGGGACAGGTC 900  
DB 939 CCAGAACAGGTTGATGCTGATGAGCATCTGACAGCTGGAGTGGGACAGGTC 998  
QY 901 ATGATGATGGCGGTGAGGCTTTATATCATGGAACACTCTCACTTATTAAGATCTT 960  
DB 999 ATGATGATGGCGGTGAGGCTTTATATCATGGAACACTCTCACTTATTAAGATCTT 1058  
QY 961 GGGCTGCTCCAAAGAGAGACTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
DB 1059 GGGCTGCTCCAAAGAGAGACTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1118  
QY 1021 GTTGTGCTCTTCAGATTTATCACTTACAGAGGTAATATTTTCAACTACAGTCTGGTG 1080  
DB 1119 GTTGTGCTCTTCAGATTTATCACTTACAGAGGTAATATTTTCAACTACAGTCTGGTG 1178



```

Db      819 ATGATGCAAGAGATGCTTCTCATGGAGTCAAAATTTGATTCACGCTAAAGATGGGGCA 878
Qy      781 AAGACTCCACGATGATGATCTCTCAACACTGAGAGAGATGATCTGGAGCAAAAT 840
Db      879 AAGACTCCACGATGATGATCTCTCAACACTGAGAGAGATGATCTGGAGCAAAAT 938
Qy      841 CCAGAACAGTGTGTCTGCTGAGATCTGACAGCTGGAGATGTTGGCAGGGTCC 900
Db      939 CCAGAACAGTGTGTCTGCTGAGATCTGACAGCTGGAGATGTTGGCAGGGTCC 998
Qy      901 ATGATGATGCGGGTGGAGCTTTATATCATGGAGACCTCTCACTATTAAAGATCTT 960
Db      999 ATGATGATGCGGGTGGAGCTTTATATCATGGAGACCTCTCACTATTAAAGATCTT 1058
Qy      961 GGGCGGCTCCAAAGAGACTGCGGGCTGCTGAGATCTGAGAGAGAGAGAGAGAGTGA 1020
Db      1059 GGGCGGCTCCAAAGAGACTGCGGGCTGCTGAGATCTGAGAGAGAGAGAGAGTGA 1118
Qy      1021 GTTGTGCTCTTCAGATATTTATCAATTAACAAGTAAATTTCCACTACAGTCTGGTG 1080
Db      1119 GTTGTGCTCTTCAGATATTTATCAATTAACAAGTAAATTTCCACTACAGTCTGGTG 1178
Qy      1081 ATGAGCTGACGCGAGAACTTTTACCCTGCGCTGCAATTCATCGGAGTGA 1140
Db      1179 ATGAGCTGACGCGAGAACTTTTACCCTGCGCTGCAATTCATCGGAGTGA 1238
Qy      1141 GCCAGGGCTCATGAGAGAGTGTGAGCTGCTGAGCCCTGCAATTAATCATCTAGGTC 1200
Db      1239 GCCAGGGCTCATGAGAGAGTGTGAGCTGCTGAGCCCTGCAATTAATCATCTAGGTC 1298
Qy      1201 CTGAGCCATGAGAGAGAGAGAGATCAACTTTTGGATTCAGCTGAGTCTGAGACC 1260
Db      1299 CTGAGCCATGAGAGAGAGAGAGATCAACTTTTGGATTCAGCTGAGTCTGAGACC 1358
Qy      1261 AGTCTACTTGAATGACTTATTAACAAGTATTTCTTCTTCATCACTCCAGAGACCAATG 1320
Db      1359 AGTCTACTTGAATGACTTATTAACAAGTATTTCTTCTTCATCACTCCAGAGACCAATG 1418
Qy      1331 ACTGTCATGATCCAAAGAGATGATGCTGCTGTTGGGCTGTTCTTAT 1380
Db      1419 ACTGTCATGATCCAAAGAGATGATGCTGCTGTTGGGCTGTTCTTAT 1478
Qy      1381 GTTGTGAGAGATGAGAAATGCTGCTAGGTC 1416
Db      1479 GTTGTGAGAGATGAGAAATGCTGCTAGGTC 1514

RESULT 10
AAZ58313
ID      AAZ58313 standard; cDNA, 1884 BP.
XX      XX
AC      AAZ58313;
XX      XX
DT      08-MAY-2000 (first entry)
XX      XX
DE      Human peptidase NALAD-ase IV cDNA.
XX      XX
KW      NALAD-ase IV, N-acetylated alpha-linked acidic dipeptidase; human;
KW      chromosome 8q21.3; prostate cancer; neurodegenerative disease;
KW      Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;
KW      peripheral neuropathy; Huntington's disease; acute brain injury;
KW      multiple sclerosis; peripheral nerve trauma; ischaemia; dementia;
KW      gene therapy; diagnosis; noctropic; neuroprotective; neuroleptic;
KW      antiparkinsonian; anticonvulsant; vasotropic; ss.
XX      XX
OS      Homo sapiens.
XX      XX
FH      Key      location/Qualifiers
FT      CDS      149..1567
FT      /tag= a
XX      XX
PN      MO200004157-A2.
XX      XX

```

```

PD      27-JAN-2000.
XX      XX
PE      14-JUL-1999; 99MO-GB02241.
XX      XX
PR      14-JUL-1999; 98GB-0015284.
XX      XX
PA      (JAN) JANSSEN PHARM NV.
XX      XX
PI      Pangalos M, Neefs JEFM, Peeters DCG;
XX      XX
DR      WPI; 2000-182424/16.
XX      XX
DR      P-PSDB; AAY58879.
XX      XX
PT      New human N-acetylated alpha-linked acidic dipeptidases for treating
PT      neural disorders e.g. Alzheimer's disease, schizophrenia and
PT      Parkinson's disease.
XX      XX
PS      Claim 7; Fig 5; 95pp; English.
XX      XX
CC      The present sequence is that of cDNA coding for human
CC      N-acetylated alpha-linked acidic dipeptidase IV (NALAD-ase IV,
CC      see AAY58879). The cDNA was obtained from a gall bladder cDNA
CC      library. Analysis of the open reading frame predicts a type II
CC      integral membrane protein with 5 potential N-glycosylation sites.
CC      The NALAD-ase II gene was mapped to chromosome 8q21.3.
CC      NALAD-ase IV expression was low in all tissues examined by RT-PCR.
CC      The invention provides human NALAD-ase I, II and IV polypeptides,
CC      cDNAs, antisense nucleic acids, vectors, host cells, transgenic
CC      organisms, antagonists and agonists. These are useful for treating
CC      neural disorders such as Alzheimer's disease, schizophrenia, ALS,
CC      Parkinson's disease, peripheral neuropathy, Huntington's disease,
CC      acute brain injury, multiple sclerosis, exposure to neurotoxic,
CC      peripheral nerve trauma, ischaemia or dementia (claimed). Nucleic
CC      acids can also be used for gene therapy and for genetic screening
CC      of predisposition to disorders associated with NALAD-ase.
XX      XX
SQ      Sequence 1884 BP; 520 A; 432 C; 458 G; 474 T; 0 other;

Query Match      99.8%; Score 1412.8; DB 21; Length 1884;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 ATGAAATTCCTTATCTTCGATTTTGGTGTGCTTACCTTTATCCCTGCTCTGG 60
Db      149 ATGAAATTCCTTATCTTCGATTTTGGTGTGCTTACCTTTATCCCTGCTCTGG 208
Qy      61 AAAGCTATATGCAAGATGCGATCTTAAGAGACCTTTGAAGAAATTAAGAAATA 120
Db      209 AAAGCTATATGCAAGATGCGATCTTAAGAGACCTTTGAAGAAATTAAGAAATA 268
Qy      121 GCCAGCTGTGAGAGATGTGTGTAAGAGAAATCATCACTAGCTGTTATGTAAGGCCAG 180
Db      269 GCCAGCTGTGAGAGATGTGTGTAAGAGAAATCATCACTAGCTGTTATGTAAGGCCAG 328
Qy      181 AACAGATCTATGAGCGATTTGCACTTCTGTTGATATCTGTGAGACCCAGACTGAGTGC 240
Db      329 AACAGATCTATGAGCGATTTGCACTTCTGTTGATATCTGTGAGACCCAGACTGAGTGC 388
Qy      241 TCAGAGACCTTAAGAAAAGCGATTCMAATTAATGTAACCAAACTGCGAGAGATGGCTG 300
Db      389 TCAGAGACCTTAAGAAAAGCGATTCMAATTAATGTAACCAAACTGCGAGAGATGGCTG 448
Qy      301 GAGAAATTTCACTGAGCGAGAGTAATCCCACTGGAGAGAGAGAGAGAGATCAGCT 360
Db      449 GAGAAATTTCACTGAGCGAGAGTAATCCCACTGGAGAGAGAGAGAGAGATCAGCT 508
Qy      361 GTGATGCTGAGAGCAAGAAATTTCAATTAAGATAGCATCTGGGCTTTGGCAGACAGATTGG 420
Db      509 GTGATGCTGAGAGCAAGAAATTTCAATTAAGATAGCATCTGGGCTTTGGCAGACAGATTGG 568
Qy      421 ACTCTTCAGAGAGCATTAAGAGAGATTTCTGGTGTGAACCTTTTCATGAATCTGAG 480
Db      569 ACTCTTCAGAGAGCATTAAGAGAGATTTCTGGTGTGAACCTTTTCATGAATCTGAG 628

```





```

Db      179 ATGAAATCCCTATCTTGCAGATTTTTCGTTGGTGTACCTTTATTCCTGCTGGCTGG 238
Qy      61 AAAGCTATATGCAAGAAATGGCATCTCTAAGAGATTTTGAAGAAATTAAGAAATA 120
Db      239 AAAGCTATATGCAAGAAATGGCATCTCTAAGAGATTTTGAAGAAATTAAGAAATA 298
Qy      121 GCCAGCTGTGAGATGTTGTCTAAAGCAATCACTAGCTGTGTATGTAAAGCCGAG 180
Db      299 GCCAGCTGTGAGATGTTGTCTAAAGCAATCACTAGCTGTGTATGTAAAGCCGAG 358
Qy      181 AACAGATCTATAGACGATGGCACTTGTGTGTATGTGAAGCCGAGCTGAGTGGC 240
Db      359 AACAGATCTATAGACGATGGCACTTGTGTGTATGTGAAGCCGAGCTGAGTGGC 418
Qy      241 TCCAGAACTTGAAGAAAGCCATCCCAATTTATGCCAAACCTGACAGCAAGATGGCTG 300
Db      419 TCCAGAACTTGAAGAAAGCCATCCCAATTTATGCCAAACCTGACAGCAAGATGGCTG 478
Qy      301 GAGAAAGTTCACTGTGAGCCAGTGAATACCCCACTGGAGAGAGGAGAAATCAGCT 360
Db      479 GAGAAAGTTCACTGTGAGCCAGTGAATACCCCACTGGAGAGAGGAGAAATCAGCT 538
Qy      361 GTGATGTGAGCCAGAAATTCATTAAGATAGCCATCTGGGTCTTGGCAGCAGATTGGG 420
Db      539 GTGATGTGAGCCAGAAATTCATTAAGATAGCCATCTGGGTCTTGGCAGCAGATTGGG 598
Qy      421 ACTCTCCAGAAAGGATTTACAGCAGAAATTTCTGTGTGTGACCTCTTTGATGAATGAG 480
Db      599 ACTCTCCAGAAAGGATTTACAGCAGAAATTTCTGTGTGTGACCTCTTTGATGAATGAG 658
Qy      481 AGAAGGGCTCTGAAGAAAGAGAGATGTTGTTTAAACCAACCTTACCTCACTAC 540
Db      659 AGAAGGGCTCTGAAGAAAGAGAGATGTTGTTTAAACCAACCTTACCTCACTAC 718
Qy      541 TCAAGAGCGGTGCAATACCGAACGAGGGGGGCGGTGAAGCTGCCAGGTGGGGCTTTG 600
Db      719 TCAAGAGCGGTGCAATACCGAACGAGGGGGGCGGTGAAGCTGCCAGGTGGGGCTTTG 778
Qy      601 GCATCTCTCATTTCCATTCGCTGGCTCTTCTTCATCTACAGTCTCAACAGGATTTAG 660
Db      779 GCATCTCTCATTTCCATTCGCTGGCTCTTCTTCATCTACAGTCTCAACAGGATTTAG 838
Qy      661 GAATACAGAGATGGGGTGGCCCAAAATTCACACAGCTGATTAAGGGTGAAGATCAGAA 720
Db      839 GAATACAGAGATGGGGTGGCCCAAAATTCACACAGCTGATTAAGGGTGAAGATCAGAA 898
Qy      721 ATGATGTCAAGAAATGCTTCTCATGAGATCAAAATTTGTCATTCAGCTTAAGATGGGCA 780
Db      899 ATGATGTCAAGAAATGCTTCTCATGAGATCAAAATTTGTCATTCAGCTTAAGATGGGCA 958
Qy      781 AAGACCTTACCAAGATCTGATTTCTTCAACACTGTAGCAGAGATCTGGGAGCAAAAT 840
Db      959 AAGACCTTACCAAGATCTGATTTCTTCAACACTGTAGCAGAGATCTGGGAGCAAAAT 1018
Qy      841 COAGAACAGTGTGTCTGTGATGATCTGACATCTGACAGCTGGGATGTTGGGAGGGTGGC 900
Db      1019 COAGAACAGTGTGTGTGATGATCTGACATCTGACAGCTGGGATGTTGGGAGGGTGGC 1078
Qy      901 ATGATGATGCGGTGAGACCTTTATATCATGGAAGACCTCTCACTTATTAAGATCTT 960
Db      1079 ATGATGATGCGGTGAGACCTTTATATCATGGAAGACCTCTCACTTATTAAGATCTT 1138
Qy      961 GGGGTGCGTCCAAAGAGACTCTGGGCTGTGCTGTGAGCTGCAAGAAACAGAGTGA 1020
Db      1139 GGGGTGCGTCCAAAGAGACTCTGGGCTGTGCTGTGAGCTGCAAGAAACAGAGTGA 1198
Qy      1021 GTTGTGCTTCCAGATATTAAGTAAACAAGATTAATTTCCAACTACAGTGTGGG 1080
Db      1199 GTTGTGCTTCCAGATATTAAGTAAACAAGATTAATTTCCAACTACAGTGTGGG 1258
Qy      1081 ATGAGTGTGACGACAGAACTTCTTACCACTGGGCTGCAATTCAGTGGCAAG 1140

```

```

Db      1259 ATGAGTGTGACGACGAGAACTTCTTACCACTGGGCTGCAATTCAGTGGCAAGAAAG 1318
Qy      1141 GCCAGGGCCATCATGAGAGAGTTATAGCTGTGAGCCCTCATATCATCAGAGTC 1200
Db      1319 GCCAGGGCCATCATGAGAGAGTTATAGCTGTGAGCCCTCATATCATCAGAGTC 1378
Qy      1201 CTGAGCCATGAGAGAGAGAGATCACTTTTGAATCCAAAGCTGAGTGGCTTGAGCC 1260
Db      1379 CTGAGCCATGAGAGAGAGAGATCACTTTTGAATCCAAAGCTGAGTGGCTTGAGCC 1438
Qy      1261 AGTCTATTGATGACTTATACAGATATTTCTTCCATCATCTCCACGAGACACATG 1320
Db      1439 AGTCTATTGATGACTTATACAGATATTTCTTCCATCATCTCCACGAGACACATG 1498
Qy      1321 ACTGTCATGATCCCAAGCAGATGAATGTGCTGCTGTTGGGCTGTTTCTTAT 1380
Db      1499 ACTGTCATGATCCCAAGCAGATGAATGTGCTGCTGTTGGGCTGTTTCTTAT 1558
Qy      1381 GTTGTGACAGACATGAGAAATGCTCCTAGTCC 1416
Db      1559 GTTGTGACAGACATGAGAAATGCTCCTAGTCC 1594

RESULT 12
ACCS0603
ID   ACCS0603 standard; cDNA; 2077 BP.
XX
AC   ACCS0603;
XX
DT   12-JUN-2003 (first entry)
XX
DE   Human secreted protein coding sequence, SEQ ID 270.
XX
KW   Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cyrostatic;
KW   vulnerary; antiinflammatory; nootropic; neuroprotective;
KW   antiparkinsonian; gene therapy; human; cardiovascular disorder;
KW   gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200295010-A2.
XX
PD   28-NOV-2002.
XX
PE   19-MAR-2002; 2002WO-US09785.
XX
PR   21-MAR-2001; 2001US-277340P.
PR   19-JUL-2001; 2001US-306171P.
PR   13-NOV-2001; 2001US-331287P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2003-129429/12.
XX
PT   Novel human secreted proteins, useful for detecting, preventing,
PT   diagnosing, prognosticating, treating and/or ameliorating
PT   cardiovascular disorders such as arrhythmia -
XX
PS   Claim 21; SEQ ID 270; 1881bp; English.
XX
CC   The present invention relates to novel human secreted proteins
CC   (ABR4/633-ABR49145) and their coding sequences (ACCS0344-ACCS0856). The
CC   proteins and their coding sequences are useful for the preparation of a
CC   diagnostic or pharmaceutical composition for diagnosing or treating a
CC   cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,
CC   coronary arteriosclerosis and myocardial ischemia), neural disorders,
CC   immune system disorders, muscular disorders, reproductive disorders,
CC   gastrointestinal disorders, pulmonary disorders, renal disorders,
CC   proliferative disorders and/or cancerous diseases and conditions, for
CC   wound healing and epithelial cell proliferation, to treat inflammation or
CC   infection, for treating thrombosis and arteriosclerosis, for treating or

```





XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 PI Rosen CA, Ruben SM;  
 XX WPI: 2003-029900/02.  
 DR P-PSDB; ABR00152.  
 XX

PT New human secreted proteins and nucleic acids, useful for detecting,  
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating,  
 PT e.g. gastrointestinal diseases and disorders, or cancers -  
 XX  
 XX

Claim 21: Page 841-842; 1216p; English.

XX AB271190-AB271478 represent cDNAs corresponding to 178 human secreted  
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.  
 CC AB271479-AB271540 represent human secreted protein genomic fragments. The  
 CC invention also encompasses antibodies specific for the secreted proteins,  
 CC the use of the secreted proteins in drug screening, and recombinant  
 CC vectors and host cells comprising a nucleic acid of the invention. The  
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
 CC fragments specific for the secreted proteins, and modulators of protein  
 CC activity are useful for diagnosing, treating, ameliorating or preventing  
 CC digestive disorders. Such conditions include disorders of the mouth,  
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary  
 CC tract and pancreas, and include cancers of these organs and tissues. The  
 CC secreted proteins and their nucleic acids may also be used in the  
 CC treatment of immune disorders, inflammation, infection,  
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids  
 CC of the invention may be used for chromosome identification, chromosome  
 CC mapping, in gene therapy, for identifying individuals from minute  
 CC biological samples, as hybridisation probes, and as molecular weight  
 CC markers. The present sequence represents a human secreted protein-  
 CC encoding cDNA clone of the invention.  
 XX

SQ Sequence 2077 BP, 597 A; 436 C; 482 G; 562 T; 0 other;

Query Match 99.8%; Score 1412.8; DB 25; Length 2077;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAATTCCTTATCTTGCATTTTTCGCTGCTTACCTTTATCCCTGCTGCGG 60  
 DB 132 ATGAATTCCTTATCTTGCATTTTTCGCTGCTTACCTTTATCCCTGCTGCGG 191  
 QY 61 AAACCTATATGCAAGATGCGATCTTAAAGAGACTTTTGAAGAAATTAAGAAATA 120  
 DB 192 AAACCTATATGCAAGATGCGATCTTAAAGAGACTTTTGAAGAAATTAAGAAATA 251  
 QY 121 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAG 180  
 DB 252 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAG 311  
 QY 181 AACGATCTTATGAGCGATGCGACTTCTGTGTTGATCTTTGAGCCAGCTAGTGGC 240  
 DB 312 AACGATCTTATGAGCGATGCGACTTCTGTGTTGATCTTTGAGCCAGCTAGTGGC 371  
 QY 241 TCCAGAACTTAAAGAAAGCCATCCAAATTTATGTAACAAACCTGACAGATGGGCTG 300  
 DB 372 TCCAGAACTTAAAGAAAGCCATCCAAATTTATGTAACAAACCTGACAGATGGGCTG 431  
 QY 301 GAGAAAGTTTCACTGAGCCAGTGAATACCCCACTGGAGAGGGGAGAAATTCAGCT 360  
 DB 432 GAGAAAGTTTCACTGAGCCAGTGAATACCCCACTGGAGAGGGGAGAAATTCAGCT 491  
 QY 361 GTGATGCTGAGCCAGAAATTCATAGATAGCAATCTCTGGGCTTTGGCAGCAGATTGGG 420  
 DB 492 GTGATGCTGAGCCAGAAATTCATAGATAGCAATCTCTGGGCTTTGGCAGCAGATTGGG 551  
 QY 421 ACTCTCCAGAAAGCATTAAGCAGAAAGTTCTGGTGGTGAACCTCTTTGATGAATCTGAG 480  
 DB 552 ACTCTCCAGAAAGCATTAAGCAGAAAGTTCTGGTGGTGAACCTCTTTGATGAATCTGAG 611

QY 481 AGAAGGCTCAGAGCAAGAGGAGATGTTGTTATTAACCACTTATCATCACTAC 540  
 DB 612 AGAAGGCTCAGAGCAAGAGGAGATGTTGTTATTAACCACTTATCATCACTAC 671  
 QY 541 TCAAGACGGTGCATATCCGAAGCGAGGGGCGGTGAGCTGCCAAGGTGGGGCTTTG 600  
 DB 672 TCAAGACGGTGCATATCCGAAGCGAGGGGCGGTGAGCTGCCAAGGTGGGGCTTTG 731  
 QY 601 GCATCTCTCAATTCGATCCGGGCTCTCTTCATCTACAGTCTCAGACAGGATTCAG 660  
 DB 732 GCATCTCTCAATTCGATCCGGGCTCTCTTCATCTACAGTCTCAGACAGGATTCAG 791  
 QY 661 GAATACCAAGATGCGTGGCCCAAAATTCACAAGCTGTATTAACGTGGAAGTGCAGAA 720  
 DB 792 GAATACCAAGATGCGTGGCCCAAAATTCACAAGCTGTATTAACGTGGAAGTGCAGAA 851  
 QY 721 ATGATGTCAGAAATGCTTCTCATGAGGATCAAAATTTCTATTAAGTAAAGTGGGCA 780  
 DB 852 ATGATGTCAGAAATGCTTCTCATGAGGATCAAAATTTCTATTAAGTAAAGTGGGCA 911  
 QY 781 AAGACCTAACCAATATCTGATTCCTTCAACATCTGACAGAGATCACTGGAGCAATAT 840  
 DB 912 AAGACCTAACCAATATCTGATTCCTTCAACATCTGACAGAGATCACTGGAGCAATAT 971  
 QY 841 CCAAGACAGGTTGATCTGATGATGACATCTGACAGCTGGAGATGTTGGGAGGAGTGC 900  
 DB 972 CCAAGACAGGTTGATCTGATGATGACATCTGACAGCTGGAGATGTTGGGAGGAGTGC 1031  
 QY 901 ATGATGATGAGCGGTGAGGCTTTTATATCAATGAGAGCACTCTCATTTAATTAAGTCTT 960  
 DB 1032 ATGATGATGAGCGGTGAGGCTTTTATATCAATGAGAGCACTCTCATTTAATTAAGTCTT 1091  
 QY 961 GGGGTGCTCCAAAGAGGACTCTGGGCTGGTGTCTGAGCTGAGAGAAACAAGTGGGA 1020  
 DB 1092 GGGGTGCTCCAAAGAGGACTCTGGGCTGGTGTCTGAGCTGAGAGAAACAAGTGGGA 1151  
 QY 1021 GTTGTGCTCTTCAAGTATTAATCACTTACACAGGTAATATTTTCAACTACAGTGGTG 1080  
 DB 1152 GTTGTGCTCTTCAAGTATTAATCACTTACACAGGTAATATTTTCAACTACAGTGGTG 1211  
 QY 1081 ATGAGTCTGACGAGAGACCTTTCTTAACCACTGGGCTGCAATTCACCTGCAAGTAAAG 1140  
 DB 1212 ATGAGTCTGACGAGAGACCTTTCTTAACCACTGGGCTGCAATTCACCTGCAAGTAAAG 1271  
 QY 1141 GCCAGGCGCATATGAGAGGAGTTATGAGCTGTGAGCCCTCAATATCACTCAAGGTC 1200  
 DB 1272 GCCAGGCGCATATGAGAGGAGTTATGAGCTGTGAGCCCTCAATATCACTCAAGGTC 1331  
 QY 1201 CTGAGCCATGAGAGAGGAGACATCACTTTTGAATCCAGCTGAGTGGCTGAGGCC 1260  
 DB 1332 CTGAGCCATGAGAGAGGAGACATCACTTTTGAATCCAGCTGAGTGGCTGAGGCC 1391  
 QY 1261 AGTCTACTGATGATCTTAATCAAGATTTCTTCTTCATCATCTCCACGAGACACCATG 1320  
 DB 1392 AGTCTACTGATGATCTTAATCAAGATTTCTTCTTCATCATCTCCACGAGACACCATG 1451  
 QY 1321 ACTGTCAATGATCCAAAGCAGATGAATGTCGCTGCTGCTTTGGGCTGTTGTTCTAT 1380  
 DB 1452 ACTGTCAATGATCCAAAGCAGATGAATGTCGCTGCTGCTTTGGGCTGTTGTTCTAT 1511  
 QY 1381 GTTGTGCAAGATGAGAAATGCTGCTAGGTCC 1416  
 DB 1512 GTTGTGCAAGATGAGAAATGCTGCTAGGTCC 1547

RESULT 14  
 AAK94491  
 ID AAK94491 standard; cdna; 1860 BP.  
 XX AAK94491;  
 AC  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX

DE Human full-length cDNA, SEQ ID NO: 3328.  
 XX Human, full length cDNA; cDNA synthesis; oligo-capping; ss.  
 XX Homo sapiens.  
 XX EPI130094-A2.  
 XX 05-SEP-2001.  
 PD 07-JUL-2000; 2000EP-0114089.  
 PF 08-JUL-1999; 99JP-0194486.  
 PR 11-JAN-2000; 2000JP-0118774.  
 PR 02-MAY-2000; 2000JP-0183765.  
 XX (HELI-) HELIX RES INST.  
 PA Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
 DR WPI: 2001-524255/58.  
 DR P-PSDB; AAM93559.  
 XX 830 Primers useful for synthesizing full length cDNA clones and their  
 PT use in genetic manipulation -  
 XX  
 PS Claim 8; SEQ ID NO 3328; 1380bp + sequence listing; English.  
 XX  
 CC The invention relates to primers for synthesizing full length cDNA  
 CC clones. 830 cDNA molecules encoding a human protein have been  
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
 CC molecules have been determined. Primers for synthesizing the full length  
 CC cDNA are useful for clarifying the function of the protein encoded by  
 CC the cDNA. The full length clones were obtained by construction of full  
 CC length enriched cDNA libraries that were synthesized by the oligo-capping  
 CC method. The primers enable the production of the full length cDNA easily  
 CC without any special methods. The present sequence is a full length  
 CC human cDNA of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in CD-ROM format directly from EPO.  
 CC  
 XX  
 XX  
 SQ Sequence 1860 BP; 515 A; 423 C; 451 G; 471 T; 0 other;  
 Query Match 99.4%; Score 1408; DB 22; Length 1860;  
 Best Local Similarity 99.6%; Pred. No. 0; Mismatches 5; Indels 0; Gaps 0;  
 Matches 1411; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 ATGAATTCCTTATCTTCGATTTTGGTGGTTCACCTTTATCCCTGCTGGG 60  
 DB 129 ATGAATTCCTTATCTTCGATTTTGGTGGTTCACCTTTATCCCTGCTGGG 188  
 QY 61 AAGCTTATGCAAGATGCGATCTTGAAGGACCTTTGAAGAATPAAAGAAATA 120  
 DB 189 AAGCTTATGCAAGATGCGATCTTGAAGGACCTTTGAAGAATPAAAGAAATA 248  
 QY 121 GCCAGCTGGAGATGTTGCTAAGCAATCACTAGCTGTTTATGTAAGCCGAG 180  
 DB 249 GCCAGCTGGAGATGTTGCTAAGCAATCACTAGCTGTTTATGTAAGCCGAG 308  
 QY 181 AACGATCTATGACGATTTGGCACTTCTGTTGATATCTTTGAGCCAGAGTGGC 240  
 DB 309 AACGATCTATGACGATTTGGCACTTCTGTTGATATCTTTGAGCCAGAGTGGC 368  
 QY 241 TCCAGAACCTAGAAAAAGCCATCCAAATTTGTAACAAAACCTGACAGAAATGGCTG 300  
 DB 369 TCCAGAACCTAGAAAAAGCCATCCAAATTTGTAACAAAACCTGACAGAAATGGCTG 428  
 QY 301 GAGAAATTCACCTGGAGCCAGTGAATACCCACTGGGAGAGGAGAAATTCAGCT 360  
 DB 429 GAGAAATTCACCTGGAGCCAGTGAATACCCACTGGGAGAGGAGAAATTCAGCT 488  
 QY 361 GTGATGCTGAGCCAGAAATTCATAGATAGCCATCTCGGCTTTGGCAGCAGCATGGG 420

DB 489 GTGATGCTGAGCCAGAAATTCATAGATAGCCATCTCGGCTTTGGCAGCAGCATGGG 548  
 QY 421 ACTCCTCAGAGGCAATTAACAGAGAGTCTGGTGGTACCTCTTTCAGTGAACCTGAG 480  
 DB 549 ACTCCTCAGAGGCAATTAACAGAGAGTCTGGTGGTACCTCTTTCAGTGAACCTGAG 608  
 QY 481 AGAAGGCTCTCAGAGGCAAGAGGAGATTTGTTTATACCAACCTTATCATCACTAC 540  
 DB 609 AGAAGGCTCTCAGAGGCAAGAGGAGATTTGTTTATACCAACCTTATCATCACTAC 668  
 QY 541 TCAAGAGCGTGAATTCGACAGCGAGGCGGTGGAGCTGCGAAGTGGGGCTTTTG 600  
 DB 669 TCAAGAGCGTGAATTCGACAGCGAGGCGGTGGAGCTGCGAAGTGGGGCTTTTG 728  
 QY 601 GCATCTCATTTGGATCCGTTGCTCTCTTCCATCTACAGTCTCACAAGGATTCAG 660  
 DB 729 GCATCTCATTTGGATCCGTTGCTCTCTTCCATCTACAGTCTCACAAGGATTCAG 788  
 QY 661 GAATACAGGATGGCGTCCCAAAATTCACAGCCTGTATTAACGTTGAGAAATGCAGAA 720  
 DB 789 GAATACAGGATGGCGTCCCAAAATTCACAGCCTGTATTAACGTTGAGAAATGCAGAA 848  
 QY 721 ATGATGTCAAGAAATGCTTCTCATGGGATCAAAATTTGTCATTGACTAAAGATGGGGCA 780  
 DB 849 ATGATGTCAAGAAATGCTTCTCATGGGATCAAAATTTGTCATTGACTAAAGATGGGGCA 908  
 QY 781 AAGACCTACCCAGATATGATTTCTTCAACACTGTACAGAGATCACTGGAGCAATAT 840  
 DB 909 AAGACCTACCCAGATATGATTTCTTCAACACTGTACAGAGATCACTGGAGCAATAT 968  
 QY 841 CCAGAACAGTTGATCTGTCAGTGCATCTGGAACAGTGGAGTGTGGGAGGGTGGCC 900  
 DB 969 CCAGAACAGTTGATCTGTCAGTGCATCTGGAACAGTGGAGTGTGGGAGGGTGGCC 1028  
 QY 901 ATGATGATGGCGGTGGAGCCTTTATATCATGGGAAGCACTTCACTTATTAAGATCTT 960  
 DB 1029 ATGATGATGGCGGTGGAGCCTTTATATCATGGGAAGCACTTCACTTATTAAGATCTT 1088  
 QY 961 GGGGTGCTCCAAAGAGAGATCTGCGGCTGTCTCTGACTGACAGAAAGAGGTGGA 1020  
 DB 1089 GGGGTGCTCCAAAGAGAGATCTGCGGCTGTCTCTGACTGACAGAAAGAGGTGGA 1148  
 QY 1021 GTTGGTCTCTCAGATATATCATGTTACACAGGTAATTTTCAACTACAGTCTGGTG 1080  
 DB 1149 GTTGGTCTCTCAGATATATCATGTTACACAGGTAATTTTCAACTACAGTCTGGTG 1208  
 QY 1081 ATGAGTCTGACGAGGAACCTTCTTACCACTGGGCTGCAATTCATCTGACAGTGAAG 1140  
 DB 1209 ATGAGTCTGACGAGGAACCTTCTTACCACTGGGCTGCAATTCATCTGACAGTGAAG 1268  
 QY 1141 GCCAGGCGCATATGAGAGGATTTAGCCTGCTGCAAGCCCTCAATATCATCTCAGTCTC 1200  
 DB 1269 GCCAGGCGCATATGAGAGGATTTAGCCTGCTGCAAGCCCTCAATATCATCTCAGTCTC 1328  
 QY 1201 CTGAGCCATGGAAGAGGACAGACATCACTTTTGAATTCAGAGTGGAGGCTGGAGCC 1260  
 DB 1329 CTGAGCCATGGAAGAGGACAGACATCACTTTTGAATTCAGAGTGGAGGCTGGAGCC 1388  
 QY 1261 AGTCTACTGATGATCTTATCAAGTATTTCTTCCATCACTCCACGAGACACCATG 1320  
 DB 1389 AGTCTACTGATGATCTTATCAAGTATTTCTTCCATCACTCCACGAGACACCATG 1448  
 QY 1321 ACTGTCAATGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1380  
 DB 1449 ACTGTCAATGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1508  
 QY 1381 GTTGTGACAGATGGAAGAAATGCTGCTAGGCTC 1416  
 DB 1509 GTTGTGACAGATGGAAGAAATGCTGCTAGGCTC 1544

RESULT 15

AAH9703  
ID AAH9703 standard; cDNA; 1895 BP.  
XX  
AC AAH9703;  
XX  
DT 16-OCT-2001 (first entry)  
XX  
DE Human protein encoding cDNA sequence SEQ ID NO:538.  
XX  
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
KW antiinflammatory; antirheumatic; antiallergic; immunosuppressive;  
KW antibacterial; endocrine; cardiant; central nervous system; virucide;  
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianemic; anaemia;  
KW antiaggregant; haemostatic; vulnery; antitumor; osteopathic; eczema;  
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytotoxic;  
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
KW antiaphyllactic; rheumatoid arthritis; septic shock; pancreatitis;  
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;  
KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
KW neurological disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200153455-A2.  
XX  
PD 26-JUL-2001.  
XX  
PF 22-DEC-2000; 2000WO-US35017.  
XX  
PR 23-DEC-1999; 99US-0471275.  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-457603/49.  
XX  
PT P-PSDB; AAM25762.  
XX  
XX Isolated human polynucleotides encoding polypeptides, useful for the  
XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection -  
XX  
XX Claim 1; Page 591; 1217bp; English.  
XX  
XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
XX AAM25963. The proteins can have activities based on the tissues and  
XX cells they are expressed in, such as: antineoplastic; antirheumatic;  
XX antiallergic; immunosuppressive; antibacterial; endocrine; cardiant;  
XX central nervous system; virucide; anti-HIV; fungicide; antimutagen;  
XX cardiovascular; antianemic; antiaggregant; haemostatic; vulnery;  
XX antitumor; osteopathic; dermatological; antiallergic; antiasthmatic;  
XX antidiabetic; cytoprotective; neuroprotective; antidepressant; nootropic;  
XX antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
XX encoding them can be used in gene therapy, antisense therapy and vaccine  
XX production. The proteins and polynucleotides are useful for screening for  
XX agonists or antagonists of a protein and for the treatment and diagnosis  
XX of disorders associated with the activity of a protein e.g. inflammation,  
XX rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
XX neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
XX infections, autoimmunity, genetic diseases, haematopoietic disorders,  
XX anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,  
XX osteoporosis, severe combined immunodeficiency, eczema, allergic  
XX rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
XX Alzheimer's disease, Parkinson's disease, neurodegenerative and  
XX neurological disorders.  
XX  
SQ Sequence 1895 BP; 530 A; 439 C; 450 G; 476 T; 0 other;

Query Match 95.7%; Score 1355.6; DB 22; Length 1895;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1407; Conservative 0; Mismatches 9; Indels 6; Gaps 4;  
QY 1 ATGAAATTCCTTATCTTCGCAATTTTGGTGTGTCACCTTTATCTCTGTGCTTGG 60  
DB 126 ATGAAATTCCTTATCTTCGCAATTTTGGTGTGTCACCTTTATCTCTGTGCTTGG 185  
QY 61 AAAGCTATATGCAAGATGGCATCTTAAGAGGACCTTGAAGAAATTAAGAAATA 120  
DB 186 AAAGCTATATGCAAGATGGCATCTTAAGAGGACCTTGAAGAAATTAAGAAATA 245  
QY 121 GCCAGCTGTGAGATGTTGTGTAAGACATCACTAGCTGTTATGTTAAAGCCAG 180  
DB 246 GCCAGCTGTGAGATGTTGTGTAAGACATCACTAGCTGTTATGTTAAAGCCAG 305  
QY 181 AACGATCTATGAGCGATTTGGCACTTCTGTTGATACCTGTGAGCCCAAGCTGAGTGGC 240  
DB 306 AACGATCTATGAGCGATTTGGCACTTCTGTTGATACCTGTGAGCCCAAGCTGAGTGGC 365  
QY 241 TCCAGAAACCTAAGAAAAGCCATCCAAATTAATGACAAAACCTGCAGCAAGAGGCTG 300  
DB 366 TCCAGAAACCTAAGAAAAGCCATCCAAATTAATGACAAAACCTGCAGCAAGAGGCTG 425  
QY 301 GAGAAAGTCACTGAGAGCCAGTGAATACCCCACTGAGAGAGGAGAGAAATCAAGCT 360  
DB 426 GAGAAAGTCACTGAGAGCCAGTGAATACCCCACTGAGAGAGGAGAGAAATCAAGCT 485  
QY 361 GTGATGTGAGAGCCAGTGAATACCCCACTGAGAGAGGAGAGAAATCAAGCT 420  
DB 486 GTGATGTGAGAGCCAGTGAATACCCCACTGAGAGAGGAGAGAAATCAAGCT 545  
QY 421 ACTCTCCAGAAAGCATTAAGCAAGAGTCTGTGTGACCTCTTCCATGAACTGAC 480  
DB 546 ACTCTCCAGAAAGCATTAAGCAAGAGTCTGTGTGACCTCTTCCATGAACTGAC 605  
QY 481 AGAAGGCTCAGAAAGCAAGAGGAAATGTTGTTATTAACCAACTTACATCACTAC 540  
DB 606 AGAAGGCTCAGAAAGCAAGAGGAAATGTTGTTATTAACCAACTTACATCACTAC 665  
QY 541 TCAAGACGCTGCAATACGAAACGAGGAGGCGGTGAGCTGCAAGGTGGGGCTTTG 600  
DB 666 TCAAGACGCTGCAATACGAAACGAGGAGGCGGTGAGCTGCAAGGTGGGGCTTTG 725  
QY 601 GCATCTCTCATTTGATTCGTCGCTCTCTCTCCATCTACAGTCTCTCAACAGATTCAG 660  
DB 726 GCATCTCTCATTTGATTCGTCGCTCTCTCTCCATCTACAGTCTCTCAACAGATTCAG 785  
QY 661 GAATACAGAGATGGCGGCCCAAAATTCGAACAGCTGTAATACGCTGGAAGATGACGAA 720  
DB 786 GAATACAGAGATGGCGGCCCAAAATTCGAACAGCTGTAATACGCTGGAAGATGACGAA 845  
QY 721 ATGATGTCAGAAATGCTTCTCATGGGATCAAAATTTGTCATTCAGCTAAAGATGGGGCA 780  
DB 846 ATGATGTCAGAAATGCTTCTCATGGGATCAAAATTTGTCATTCAGCTAAAGATGGGGCA 905  
QY 781 AAGACCTACCAATATGATTTCTTCAACACTGTACAGATCACTGGGAGCAATAT 840  
DB 906 AAGACCTACCAATATGATTTCTTCAACACTGTACAGATCACTGGGAGCAATAT 965  
QY 841 CCAGAACAGGTTGATCTGTCAGTGACATCTGGAACAGCTGGAGTGTGGGAGGGTGGC 900  
DB 966 CCAGAACAGGTTGATCTGTCAGTGACATCTGGAACAGCTGGAGTGTGGGAGGGTGGC 1025  
QY 901 ATGATGATGCGGTGAGGCTTATATCACTGGAAGCACTTCACTTATTAAGATCTT 960  
DB 1026 ATGATGATGCGGTGAGGCTTATATCACTGGAAGCACTTCACTTATTAAGATCTT 1085  
QY 961 GGGCTGGTCCAAAGAGACTCTGCGGCTGTGCTGACCTGCAAGAAACAAGTGGGA 1020  
DB 1086 GGGCTGGTCCAAAGAGACTCTGCGGCTGTGCTGACCTGCAAGAAACAAGTGGGA 1145  
QY 1021 GTTGTGCTCTTCAAGTATTTATCAAGTTAAATTTTCCAACTACAGTGTG 1080

```
Db      ||||||| 1146 GTTGTGCTTCAGTATTACAGTACAAAGTAATATTCCAACTACAGTCTG 1205
Qy      ||||||| 1081 ATGAGTCTGACGAGGAACCTTCTTACCACTGGGCTGCAATTCACTGGAGTGAAG 1140
Db      ||||||| 1206 ATGAGTCTGACGAGGAACCTTCTTACCACTGGGCTGCAATTCACTGGAGTGAAG 1265
Qy      ||||||| 1141 GCCAGGCAATCATGAGAGGTTATGAGCTGCTGCAGCCCTCAATATCATCTAGGTC 1200
Db      ||||||| 1266 GCCAGGCAATCATGAGAGGTTATGAGCTGCTGCAGCCCTCAATATCATCTAGGTC 1325
Qy      ||||||| 1201 CTGAGCCATGAGAGGAGACACATCACTTTTGATCCAACTGAGTGCTGAGACC 1260
Db      ||||||| 1326 CTGAGCCATGAGAGGAGACACATCACTTTTGATCCAACTGAGTGCTGAGACC 1385
Qy      ||||||| 1261 AGTCTACTGATGACTTATTAAGATTTCTTCCATCACTCCACGAGACACCATG 1320
Db      ||||||| 1386 AGTCTACTGATGACTTATTAAGATTTCTTCCATCACTCCACGAGACACCATG 1445
Qy      ||||||| 1321 ACTGTC--ATGATCCAAAGCAGATGATG-TTGCTGCTGCTG-TTGAGGCTGTTGTT 1375
Db      ||||||| 1446 ACTGTCATGGGATCCAAACGCAATGATGTTGCTGCTGCTGTTTGGGCTGTTGTT 1505
Qy      ||||||| 1376 CTTATGT-TGTTGACAGACATGGAAGAAATGCTGCTAGGTCC 1416
Db      ||||||| 1506 CTTATGTGTGACAGACATGGAAGAAATGCTGCTAGGTCC 1547
```

Search completed: December 22, 2003, 15:10:01  
Job time : 346.284 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:22:34 ; Search time 2771.56 Seconds  
(without alignments)  
12417.225 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_99\_1514

Perfect score: 1416  
Sequence: 1 ATGAAATTCCTTACTTCG.....AGGAATGCTGCTAGGTCC 1416

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:  
1: em\_estba:\*\*  
2: em\_estba:\*\*  
3: em\_estba:\*\*  
4: em\_estba:\*\*  
5: em\_estba:\*\*  
6: em\_estba:\*\*  
7: em\_estba:\*\*  
8: em\_estba:\*\*  
9: gb\_est1:\*\*  
10: gb\_est2:\*\*  
11: gb\_est3:\*\*  
12: gb\_est4:\*\*  
13: gb\_est5:\*\*  
14: gb\_est6:\*\*  
15: em\_estfun:\*\*  
16: em\_estfun:\*\*  
17: em\_gss\_hum:\*\*  
18: em\_gss\_hum:\*\*  
19: em\_gss\_hum:\*\*  
20: em\_gss\_hum:\*\*  
21: em\_gss\_hum:\*\*  
22: em\_gss\_hum:\*\*  
23: em\_gss\_hum:\*\*  
24: em\_gss\_hum:\*\*  
25: em\_gss\_hum:\*\*  
26: em\_gss\_hum:\*\*  
27: em\_gss\_hum:\*\*  
28: gb\_gss1:\*\*  
29: gb\_gss2:\*\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Query	Match	Length	ID	Description
No.						
1	1118.8	79.0	1868	11	BC010977	Mus muscu
2	1116.8	78.9	1739	11	AK032972	Mus muscu
3	1115.2	78.8	1805	11	AK075686	Mus muscu
4	958.8	67.7	1024	11	BC012019	Homo sapi

5	905.8	64.0	1050	13	BX439467	BX439467
6	843.8	59.6	1088	13	BX355940	BX355940
7	796.8	56.3	909	13	BQ878966	AGENCOURT
8	774.4	54.7	1126	13	BX360507	BX360507
9	766.2	54.1	919	13	BX335995	BX335995
10	748.6	52.9	866	13	BX146905	AGENCOURT
11	742.8	52.5	1012	13	BX416896	BX416896
12	740	52.3	902	13	BG751497	BG751497
13	736.6	52.0	998	13	BX460463	BX460463
14	726.6	51.3	996	13	BX355939	BX355939
15	719.2	50.8	739	13	CB958693	CB958693
16	683.6	48.3	813	14	CD519131	AGENCOURT
17	683.4	48.3	922	12	BG761741	BG761741
18	681.8	48.1	708	14	CD366462	UT-H-FTL-
19	676.6	47.8	945	12	B1909780	603070868
20	666.8	47.1	1081	13	BQ072892	AGENCOURT
21	665.2	47.0	889	10	BE906771	601498329
22	663.6	46.9	801	12	B1754143	603025609
23	655.8	46.3	701	12	B1771109	603059686
24	653.2	46.1	1020	13	BX439466	BX439466
25	642.4	45.4	786	10	BG430966	602500267
26	641.8	45.3	783	12	B1754468	603022607
27	636.8	45.0	653	10	BE218907	hV46B06.X
28	632.8	44.7	822	12	B1911772	603065377
29	630.4	44.5	650	10	BG432644	602496315
30	627.8	44.3	3151	11	BC017373	Mus muscu
31	626.2	44.2	804	14	CD352906	UT-M-GLO-
32	626.2	44.2	879	13	BQ218838	AGENCOURT
33	625.6	44.2	677	10	BE670671	7638406.X
34	620	43.8	944	12	B1819273	603037713
35	614.8	43.4	1073	13	BX360506	BX360506
36	604	42.7	926	13	BQ938234	AGENCOURT
37	599.6	42.3	805	10	BG563740	602584524
38	592.6	41.9	1201	13	BX379674	BX379674
39	591.6	41.8	849	10	BG166321	602345529
40	591.2	41.8	1003	13	BX416895	BX416895
41	590.8	41.7	751	14	CD558931	AGENCOURT
42	590.8	41.7	806	12	B1764052	603043258
43	590.2	41.7	727	12	BG765433	602738911
44	583.8	41.2	736	12	B1769575	603054947
45	577.8	40.8	745	12	B1462584	603203005

#### ALIGNMENTS

RESULT 1  
LOCUS BC010977 1868 bp mRNA linear HTC 04-MAR-2003  
DEFINITION Mus musculus, similar to plasma glutamate carboxypeptidase, clone  
IMAGE:4018236, mRNA.  
ACCESSION BC010977 GI:15012137  
VERSION BC010977.1  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORANISM  
REFERENCE  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (23-JUL-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>

contact: amadan@systembiology.org  
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLM at: <http://image.llnl.gov>  
 Series: IRAX Plate: 18 Row: 9 Column: 4  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9055233  
 This clone has the following problem: retained intron.

## FEATURES

source

location/Qualifiers

1..1868  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6 J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4018296"  
 /tissue\_type="mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."  
 /clone\_1ib="NCI CGAP\_Lu29"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 BASE COUNT 505 a 418 c 460 g 485 t  
 ORIGIN

Query Match 79.0%; Score 1118.8; DB 11; Length 1868;

Best Local Similarity 87.7%; Pred. No. 0; Mismatches 172; Indels 0; Gaps 0;

Matches 1222; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

23 TTTTGGTGTGTTGACCTTTATCCCTGCTGTGGGAAAGCTATATGCAAGATGCA 82  
 217 TTCCGTTTCATGTTACCTTTTAAAGCTTGTGTAAGCTGATTAAGAAATGAG 276  
 83 TCTCTAAGAGACTTTTGAAGAAATTAAGAAATGCAAGCTGTGAGATGTTGCTA 142  
 277 TTTTCACAGCAATTTGCAAGAAATTAAGAAATGCAAGCTGTGAGATGTTGCTA 336  
 143 AAGCAATCATCAAGCTGTGTTTATGTTAAGCCCAAGAGATCTATGAGCGATTTG 202  
 337 AAGCAATCATCAAGCTGTGTTTATGTTAAGCCCAAGAGATCTATGAGCGATTTG 396  
 203 CACTTGTGTTGATCTGTTGAGACCAAGCTGAGTGTCTCAAGAACTTAAAGGCA 262  
 397 GACTTGTGTTGATCTGTTGAGACCAAGCTGAGTGTCTCAAGAACTTAAAGGCA 456  
 263 TCCAAATTAATGATCAAAACCTGCAAGAGATGAGGCTGGAAGATTTCACTGAGCA 322  
 457 TCCAAATTAATGATCAAAACCTGCAAGAGATGAGGCTGGAAGATTTCACTGAGCA 516  
 323 TGAGAAATCCCACTGGGAGAGAGAGAAATGATGCTGATGAGCGCAAGATTTG 382  
 517 TCAGAAATCCCACTGGGAGAGAGAGAAATGATGCTGATGAGCGCTCAATTC 576  
 383 ATAGATAGCCATCTGAGTCTTGGCAGAGATTTGGATCTCTCCAGAGGCAATTAC 442  
 577 ACAATATGATATTTAGAGTCTTGGCAGAGATTTGGATCTCTCCAGAGGCAATTAC 636  
 443 CAGAAATTTGAGTGTGATCTTTTGATGAATCTGCAAGAGAGGCTCTCAAGCAAG 502  
 637 CAGAAATTTGAGTGTGATCTTTTGATGAATCTTCAAGAGAGGCTCTCAAGCAAG 696  
 503 GGAATATTTGTTTATTAACCACTTACATCAATCAAGAGAGGCTCTCAAGCAAG 562  
 697 GGAATATTTGTTTATTAACCACTTACATCAATCAAGAGAGGCTCTCAAGCAAG 756  
 563 CGCAGAGGAGGCTGTGAAGAGCTGCAAGAGAGGCTTTGGCATCTCTCAATTCAGTGG 622  
 757 TGCAAGGAGGCTGTGAAGAGCTGCAAGAGAGGCTGTGCAATCTCTCAATTCAGTGG 816  
 623 CTTCTTTCTCACTCAAGATCTCTCAAGAGATTTCAAGAAATCAAGAGAGGCTGCA 682  
 817 CTTCTTTCTCACTCAAGATCTCTCAAGAGATTTCAAGAAATCAAGAGAGGCTGCA 876

683 AATTTCAACAGCTGTATTCAGGTGAGAGATGCAAGAAATGATGCAAGATGCTTCTC 742  
 877 AGATTTCACAGCCTGTATTCAGGTGAGAGATGCAAGAAATGATGCAAGATGCTTCTC 936  
 743 ATGGATTCACAAATTTGATTCATTCAGTAAAGATGGGGGCAAGAGCTTACCAATCTGATT 802  
 937 GTGGGAAACAAATTTGATTCATTCAGTAAAGATGGGGGCAAGAGCTTACCAATCTGATT 996  
 803 CCTTCAACAGTGTGAGAGATTCAGTAAAGATGGGGGCAAGAGCTTACCAATCTGATT 862  
 997 CCTTCAACAGTGTGAGAGATTCAGTAAAGATGGGGGCAAGAGCTTACCAATCTGATT 1056  
 863 GTGACATCTGAGAGAGCTGTGATTTGGGAGGCTGTCATGATGATGAGGAGGCT 922  
 1057 GTGACATCTGAGAGAGCTGTGATTTGGGAGGCTGTCATGATGATGAGGAGGCT 1116  
 923 TTATATCATGGGAGAGAGCTGTGATTTGGGAGGCTGTCATGATGATGAGGAGGCT 982  
 1117 TCATATCATGGGAGAGAGCTGTGATTTGGGAGGCTGTCATGATGATGAGGAGGCT 1176  
 983 TGCGCTGTGCTGTGAG 1042  
 1177 TGCGCTGTGCTGTGAG 1236  
 1043 AGTACACAGGTAATATTTCCAACTAGTGTGATGAGAGAGAGAGAGAGAGAGAGAG 1102  
 1237 AGTACACAGGTAATATTTCCAACTAGTGTGATGAGAGAGAGAGAGAGAGAGAGAG 1296  
 1103 TCTTACCACTGTGAG 1162  
 1297 TCTTACCACTGTGAG 1356  
 1163 TTTAGAGCTGTGAG 1222  
 1357 TCAATGATCTGTGAG 1416  
 1223 ACATCACTTTTGTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1282  
 1417 ATATTAATCTGTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1476  
 1283 AGATATTTCTTTCATCACTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1342  
 1477 AGATATTTCTTTCATCACTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1536  
 1343 TGAATTTGCTGTGAG 1402  
 1537 TGAATTTGCTGTGAG 1596  
 1403 TGCTGCTAGGCTC 1416  
 1597 TGTGCTCAGGCTC 1610

DB

RESULT 2  
 AK032972  
 LOCUS  
 DEFINITION  
 Mus musculus 12 days embryo male wolffian duct includes surrounding region cDNA, RIKEN full-length enriched library, clone:8720483C06 product:Plasma glutamate carboxypeptidase, full insert sequence.  
 ACCESSION  
 AK032972.1 GI:26328732  
 VERSION  
 AK032972.1  
 KEYWORDS  
 HTC; CAP tripper.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Mech. Enzymol. 303, 19-44 (1999)  
 JOURNAL  
 MEDLINE  
 PUBMED  
 99279253  
 10349636



REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
20499374  
11042159  
3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, O., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsuki, J., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
20530913  
11076861  
4  
Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadori, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Knehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schiraldi, L. M., Stabul, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barin, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, C., Hume, D. A., Kamita, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mommaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyono, K., Wang, K. H., Weitz, C., Whiteaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kontseki, S. and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
21085660  
11217851  
5  
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 1739)  
Aachidi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirokane, T., Horii, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kawai, H., Kawai, J., Kojima, Y., Kondo, S., Komno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
location/Qualifiers

FEATURES  
SOURCE  
1. 1739  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:6720483C06"  
/db\_xref="taxon:10090"  
/clone="6720483C06"  
/sex="male"  
/tissue\_type="wolfian duct includes surrounding region"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="12 days embryo"  
81. 1493  
/note="unnamed protein product; plasma glutamate carboxypeptidase (MGD) [MG1:1889205, GB|NW\_018755, evidence: BLASTN, 99%, match=943]  
/putative"  
/protein\_id="BAC28105.1"  
/db\_xref="GI:2632873"  
/translation="MRSFLFLVHLAUGSGAVRNGVSGTFRREIKELNEDV AKAIINAVYGRKONSVEELGLIVTGVRSLGSKNLEVAIDIMVONQDLENVH LEOVRIPHMERGESAVMEPRIRHNAIILGSGISGIPPGIAEVLVAVSPDLQR ASARGKILVYNOPTGYEVQYRQGAVALKGAVALIOSVAFSISPHGTIO KYDGGPKIPFATITVEDAMSRMNRGKIVIHLEMAKTYPDSDPNTVEITGS MYREVLVUSQTEILDSWDNGALDGGGATISBALSIVYDLGLRPRTIRLYMTAE EGGTGAQSYTEILHRANISRTSLVWADSTFLPTGQTFGSDKAPRIMEVNLQAP LNTVAFNSGEGDINFWIQAGVPSLRDLVYKFFHSHSDTWTWDPKOMVVA AVAAVAVVAVVADMDLPR8"  
1722. 1727  
/note="putative"  
1739  
/note="putative"  
BASE COUNT 482 a 369 c 423 g 465 t  
ORIGIN  
polya\_signal  
polya\_site  
Query Match 78.9%; Score 1116.8; DB 11; Length 1739;  
Best Local Similarity 87.9%; Pred. No. 0;  
Matches 1217; Conservative 0; Mismatches 167; Indels 0; Gaps 0;  
33 TGTTCACCTTTTATCCCTGTGCTCTGGGAAGCTATATGCAAGATGCATCTCTAAGAG 92  
107 TGTTCACCTTTTATCCCTGTGCTCTGGGAAGCTATATGCAAGATGCATCTCTAAGAG 166  
93 GACTTTGAAGAAATTAAGAAAGAAATAGCCAGCTGTGAGATGTTCCTAAAGCATCAT 152  
167 AACCTTTCAGAAATAAAGAAATAGCCAGCTATGAGATGTTCCTAAAGCATCAT 226  
153 CAACCTGCTTTATATGTAAAGCCAGAACAGATCTATAGAGGATTTGGCATCTTCTGCT 212  
227 CAACCTGCTTTATATGTAAAGCCAGAACAGCTATGAGATGTTCCTAAAGCATCAT 286  
213 TGATAGCTTTGAGCCAGACTGAGTGCCTCAAGAACTTGAAGAAAGCCATCCCAATTTAT 272  
287 TGATAGCTTTGAGCCAGACTGAGTGCCTTGAAGAACTTGAAGAAAGCCATCCCAATTTAT 346  
273 GTACCAAACTGAGCAAGATGGCTGAGAAAGTTCACTTGAGGCACTGAGCAATATAC 332  
347 GTACCAAACTGAGCAAGATGGCTGAGAAAGTTCACTTGAGGCACTGAGCAATATAC 406  
333 CCATGAGGAAGAGGAGAAATATCAGCTGTGATCTTGGACCAAGATTTGATAAGATAGC 392  
407 CCATGAGGAAGAGGAGAAATATCAGCTGTGATCTTGGACCTTGAATTTACAAAGATGAC 466  
393 CATCTGGGCTTGGAGCAGCATTTGGGATCTCTCCAGAGGATTTACAGCAAGATTTCT 452

Db 467 TATCTAGTCTTGGCAGCAGCATTTGGAGCTCCCCAGAGGACATCACAGCAAGTGTCT 526  
 Oy 453 GGTGTGACCTCTTTCGATGAACTCCAGAGAGGGCCCTCAAGAGAAAGGAAGATGT 512  
 Db 527 GGTGTGACCTCTTTCGATGAACTCCAGAGAGGGCCCTCAAGAGAAAGGAAGATGT 586  
 Oy 513 TGTATTAACCAACCTTATCATCACTACTCAAGAGCGGTGCATACCAAGCGAGGGGCG 572  
 Db 587 TGTATTAACCAAGCTTATCATCTGCTATGAGAGAGCTGTGCACTGCCGGGTGCAGGAGC 646  
 Oy 573 GGTGAGACCTCCCAAGGTGGGGGCTTTGGCATCTCTCATTTGCATCCGTCCCTCTTC 632  
 Db 647 TGTGGAAGCTCCCAAGGTGGGGGCTTTGGCATCTCTCATTTGCATCCGTCCCTCTTC 706  
 Oy 633 CATCTACAGTCTCTCACAGAGATTCAGAGATTCAGAGATGGCCGCCCAAAATTCAC 692  
 Db 707 CATCTACAGTCTCTCACAGAGATTCAGAGATTCAGAGATGGCCGCCCAAAATTCAC 766  
 Oy 693 AGCCTGATTTACGAGTGAAGATGCAAGATGATGTCAGAGATGGCTCTCATGGGATCAA 752  
 Db 767 AGCCTGATTTACGAGTGAAGATGCAAGATGATGTCAGAGATGGCTCTCATGGGATCAA 826  
 Oy 753 AATTGTCAATCACTGTAAGATGGGGGCAAGACCTACCCAGATCTGATCTCTTCAACAC 812  
 Db 827 AATTGTCAATCACTGTAAGATGGGGGCAAGACCTACCCAGATCTGATCTCTTCAACAC 886  
 Oy 813 TGTAGCAAGATCATCTGGGAGCAATATTCAGAGATGATGTCAGAGATGGCTCTCATGG 872  
 Db 887 TGTGTGCAAGATCATCTGGGAGCAATATTCAGAGATGATGTCAGAGATGGCTCTCATGG 946  
 Oy 873 GGAACAGTGGGATGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 932  
 Db 947 GGAACAGTGGGATGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1006  
 Oy 933 GGAACAGTCTCTCACTTATTAAGATCTTGGGCTGCTGCTCAAGAGAGGAGGAGGAGG 992  
 Db 1007 GGAACAGTCTCTCACTTATTAAGATCTTGGGCTGCTGCTCAAGAGAGGAGGAGGAGG 1066  
 Oy 993 GCTCTGAGCTGCAAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1052  
 Db 1067 GCTCTGAGCTGCAAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1126  
 Oy 1053 GGTATATATTTTCACTACAGTCTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1112  
 Db 1127 GGTATATATTTTCACTACAGTCTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1186  
 Oy 1113 TGGGCTGCAATTCAGTGGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1172  
 Db 1187 TGGGCTGCAATTCAGTGGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1246  
 Oy 1173 GGTGAGAGGCTCTCATATTCATCTCAGGCTCTGAGCCATGAGAGAGGAGGAGGAGGAGG 1232  
 Db 1247 TGTGAGAGGCTCTCATATTCATCTCAGGCTCTGAGCCATGAGAGAGGAGGAGGAGGAGG 1306  
 Oy 1233 TTTGATCAAGCTGAGTGGGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1292  
 Db 1307 CTGATTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1366  
 Oy 1293 CTTCATCATCTCCACGAGAGACCATGATCTGTCTCATGATTCAGAGAGGAGGAGGAGG 1352  
 Db 1367 TTTTCATCATCTCCACGAGAGACCATGATCTGTCTCATGATTCAGAGAGGAGGAGGAGG 1426  
 Oy 1353 TGTCTCTGTTGGGCTGTTGTTCTTATGTTGTTGATGATGATGATGATGATGATGATGAT 1412  
 Db 1427 TGTCTCTGTTGGGCTGTTGTTCTTATGTTGTTGATGATGATGATGATGATGATGATGAT 1486  
 Oy 1413 GTCC 1416  
 Db 1487 GTCC 1490

RESULT 3  
 AK075686

LOCUS AK075686 1805 bp mRNA linear HTC 07-DEC-2002  
 DEFINITION Mus musculus 18-day embryo whole body cDNA, RIKEN full-length  
 enriched library, clone:1190003p2 product:plasma glutamate  
 carboxypeptidase, full insert sequence.  
 ACCESSION AK075686  
 VERSION AK075686.1 GI:26344480  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 REFERENCES  
 AUTHORS Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 2  
 REFERENCES  
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20493374  
 PUBMED 11042159  
 3  
 REFERENCES  
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,  
 Komno, H., Akiyama, J., Nishi, K., Kizunai, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
 Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Obara, E., Watanabe, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861  
 4  
 REFERENCES  
 AUTHORS Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
 Arakawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S.,  
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamahata, I.,  
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
 Fleischmann, W., Gaasterland, T., Gisi, C., King, C., Kishida, H.,  
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Peckol, G., Quackenbush, J.,  
 Schmitt, L.M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T.,  
 Sakai, K., Okido, T., Furuno, M., Aono, H., Baidarelli, R., Barab, G.,  
 Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F.,  
 Brownstein, M.J., Bult, C., Fletcher, P., Fujita, M., Gariboldi, M.,  
 Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,  
 Marchionni, L., Mashima, J., Mazzarelli, D., Mombauts, P., Nordone, P.,  
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K.,  
 Schonbach, C., Seta, T., Shibata, Y., Schorch, K.F., Suzuki, H., Toyo-oka, K.,  
 Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wyshew-Boris, A.,  
 Yoshida, K., Hasegawa, Y., Kawaji, H., Kohzuki, S., and Hayashizaki, Y.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409 (6821), 685-690 (2001)  
 MEDLINE 21085660  
 PUBMED 11217851  
 5  
 REFERENCES  
 AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research  
 Group Phase I & II Team.  
 TITLE Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 JOURNAL Nature 420, 563-573 (2002)  
 MEDLINE 1217851  
 PUBMED 1217851  
 6 (bases 1 to 1805)  
 REFERENCES  
 AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
 Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,  
 Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F.,  
 Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,

TITLE  
JOURNAL

Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Yoshida, Y., Toyota, T., Yamamura, T., Yamane, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-APR-2002) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
URL: http://location.qualifiers

FEATURES  
source

1. 1805  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:1190003P12"  
/db\_xref="MGI:1909852"  
/db\_xref="taxon:10090"  
/clone="1190003P12"  
/issue\_type="whole body"  
/clone\_id="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="18-day embryo"  
147. 1559  
/note="unnamed protein product; plasma glutamate carboxypeptidase (MGP)|MGI:1889205, GI|NM\_018755, evidence: BLASTN, 98%, match=943)  
putative"  
/codon\_start=1  
/protein\_id="BAC35891.1"  
/db\_xref="GI:26344481"  
/db\_xref="MGI:1923585"  
/db\_xref="MGI:1923585"  
/translation="MRSFLFLIVHLALSGKAVPKXGVSORTFREIKELIANEDVAKALINLAVGKYNRSYERLGLVIVGPRLSGSKLAKAIQIMYONLQDDLENVHLEQVRIPIHMERGESAVMLEPRIRKMAILGLSSIGTPPGITAEVIVVASFBLQRLASEARGLIYVNOPTGYEKTVOYVOGAFAAKAVASLIQSVASFSIYSPHTGQKYODGVKPIPTACTVDAAEMMSMRSKNGKIVHLEMGAKTVPDTSFMTVAITGS MYREVVIVSGHLDSDVYGGALDDGGAPISMEALSLVYKDLGRPKRTKLVLTATBEQGGIGASQYIEHLKANKISKYSLVWEADSGFLPTGLQFTGSDPARAKIMEVNLQPL LNVTKVSNSEGTIDINFIQAGVGAASLRDLYKRFPHSHSGDTMTVMPEKONVAA AVMAVAVVAVADMDEMLPRS"  
1788. 1793  
/note="putative"  
1805  
/note="putative"

## CDS

polya\_signal  
1805  
/note="putative"  
polya\_site  
1805  
/note="putative"  
BASE COUNT 486 a 397 c 445 g 476 t 1 others  
ORIGIN

Query Match 78.8%; Score 1115.2; DB 11; Length 1805;  
Best Local Similarity 87.9%; Pred. No. 0;  
Matches 1216; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

Qy 33 TGTTCACCTTTTATCCCTGCTCTGGGAAGCTATATGCAAGATGCACTCTTAAGAG 92  
Db 173 TGTTCACCTTTTATCCCTGCTCTGGGAAGCTATATGCAAGATGCTTCTCAAGC 232  
Qy 93 GACTTTGAAGAAATTAAGAAATATGCAAGCTGTGGAAGTGTGCTTAAGCAATCAT 152  
Db 233 AACATTCGAGAAATTAAGAAATATGCAAGCTATGATGTTGCTTAAGCAATCAT 292

Qy 153 CAACCTAGCTGTTTATGTTAAAGCCCAAGACAGATCCTATGAGCGATGCGACTTCGT 212  
Db 293 CAACCTAGCTGTTTATGTTAAATACCAAGACCGGTCTATGAGCGTTTGGACATTCGT 352  
Qy 213 TGATACCTGTGGACCCAGACTGAGTGGCTCCAGAACTTGAAGAAAGCCATCCAAATAT 272  
Db 353 TGATACCTGTGGACCCAGACTGAGTGGCTCCAGAACTTGAAGAAAGCCATCCAAATAT 412  
Qy 273 GTACCAAAACCTGAGCAAGATGAGGCTGGAGAAAGTTCACTGAGCAGTGAATATAC 332  
Db 413 GTACCAAAACCTGAGCAAGATGAGGCTGGAGAAAGTTCACTGAGCAGTGAATATAC 472  
Qy 333 CCACCTGGAGAGGAGAGAAATCAGCTGTGATGCTGAGCCCAAGATTTATAGATAGC 392  
Db 473 CCACCTGGAGAGGAGAGAAATCAGCTGTGATGCTGAGCCCAAGATTTATAGATAGC 532  
Qy 393 CATCTGGGCTTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCATTAACAGAGATTC 452  
Db 533 TATTTGAGTCTTTGGCAGCAGCATTTGGGACTCTCCAGAGAGCATTAACAGAGATTC 592  
Qy 453 GGTGGTGAACCTTTTGGATGATGAGAGAGGAGGCTCCAGAAAGAGAGAGATTC 512  
Db 593 GGTGGTGAACCTTTTGGATGATGAGAGAGGAGGCTCCAGAAAGAGAGAGATTC 652  
Qy 513 TGTTTATACCAACCTTACATCACTCAAGAGCGTCAATACCGAAGCGAGGAGG 572  
Db 653 TGTTTATACCAACCTTACATCACTCAAGAGCGTCAATACCGAAGCGAGGAGG 712  
Qy 573 GGTGGAGCTGCGAAGGTGGGAGGCTTTGGCATCTCTCATTCGATCCGTGAGCTCTTC 632  
Db 713 TGTGGAGCTGCGAAGGTGGGAGGCTTTGGCATCTCTCATTCGATCCGTGAGCTCTTC 772  
Qy 633 CATCTAGACCTTCACACAGATTTACAGAAATCCAGAGGAGGCTGCGCAAAATTCAC 692  
Db 773 CATCTAGACCTTCACACAGATTTACAGAAATCCAGAGGAGGCTGCGCAAAATTCAC 832  
Qy 693 AGCCTGATTAACGCTGGAGATGCAAGATGATGATCAAGATGAGCTTCATGAGATCAA 752  
Db 833 AGCCTGATTAACGCTGGAGATGCAAGATGATGATCAAGATGAGCTTCATGAGATCAA 892  
Qy 753 AATTGCTATTCAGTAAAGATGGGGGCAAGAGCTTCCAGATTCGATTCCTTCAACAC 812  
Db 893 AATTGCTATTCAGTAAAGATGGGGGCAAGAGCTTCCAGATTCGATTCCTTCAACAC 952  
Qy 813 TGTAGCAGATCACTGAGGAGCAATTCAGAAACAGTGTACTGTCACTGAGTCACT 872  
Db 953 TGTAGCAGATCACTGAGGAGCAATTCAGAAACAGTGTACTGTCACTGAGTCACT 1012  
Qy 873 GGAAGCTGGAGATGTTGGGAGGAGGCTCATGATGATGAGGAGGAGCTTTATATCATG 932  
Db 1013 GGAAGCTGGAGATGTTGGGAGGAGGCTCATGATGATGAGGAGGAGCTTTATATCATG 1072  
Qy 933 GGAAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGCACTGCGGCTGT 992  
Db 1073 GGAAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGCACTGCGGCTGT 1132  
Qy 993 GCTCTGACCTGCAAGAAACAGGTGAGTGTGGCTCTCCAGATTTATGAGTACATAA 1052  
Db 1133 GCTCTGACCTGCAAGAAACAGGTGAGTGTGGCTCTCCAGATTTATGAGTACATAA 1192  
Qy 1053 GGTAAATATTTTCAACACTGAGTGTGATGAGTGTGAGGAGGAGGAGGAGGAGGAGG 1112  
Db 1193 GGTAAATATTTTCAACACTGAGTGTGATGAGTGTGAGGAGGAGGAGGAGGAGGAGG 1252  
Qy 1113 TGGGCTCAATTCACCTGAGTGAAGAAAGGCGAGGAGGAGGAGGAGGAGGAGGAGG 1172  
Db 1253 TGGGCTCAATTCACCTGAGTGAAGAAAGGCGAGGAGGAGGAGGAGGAGGAGGAGG 1312  
Qy 1173 GCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1232  
Db 1313 TCTGCAACCCCTCAATTCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1372  
Qy 1233 TTGATCCAGAGCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1292

```

Db      1373  CTGATCCAGCTGAGTACCTGAGCCAGCTGCTGATGATCTGTACAAGTATTTCTT 1432
Oy      1293  CTTGCATACCTCCACCGAGACACATGACTGTATGATCCAAAGCATGATGATGTC 1352
Db      1433  TTTCCATCATTTCCACGAGACACATGATCTGTATCCAAAGCATGATGATGTC 1492
Oy      1353  TGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGACATGAGAAATGCTGCTAG 1412
Db      1493  TGCTGCTGATGAGGCTGTTGTTGCTTATGTTGTCACATGATGATGAAATGCTGCCAG 1552
Oy      1413  GTCC 1416
Db      1553  GTCC 1556

RESULT 4
BC012019      1024 bp mRNA linear HTC 04-MAR-2003
LOCUS      Homo sapiens, similar to plasma glutamate carboxypeptidase, clone
DEFINITION      IMAGE:4455631, mRNA.
ACCESSION      BC012019.1 GI:15080558
VERSION      BC012019.1 GI:15080558
KEYWORDS      HTC.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE      1 (bases 1 to 1024)
AUTHORS      Strausberg, R.
TITLE      Direct Submision
JOURNAL      Submitted (30-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnarathne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Louised, H.,
Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LIML at: http://image.liml.gov
Series: IRAX Plate: 28 Row: b Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5174626
This clone has the following problem: retained intron.
FEATURES
source
location/Qualifiers
1..1024
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4455631"
/tissue_type="kidney, hypernephroma"
/clone_id="NIH_MGC_89"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
BASE COUNT      317 a      210 c      256 g      241 t
ORIGIN
Query Match      67.7%; Score 958.8; DB 11; Length 1024;
Best Local Similarity 99.8%; Pred. No. 1.4e-269;
Matches 960; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Oy      1  ATGAATTCCTTATCTTGCATTTTGGTGTGTGTCACCTTTATCTGTGCTGG 60
Db      36  ATGAATTCCTTATCTTGCATTTTGGTGTGTGTCACCTTTATCTGTGCTGG 95
Oy      61  AAAGCTATATGCAAGAAATGGCATCTTAAGAGACCTTTGAGAAATPAAAGAAATA 120
Db      96  AAAGCTATATGCAAGAAATGGCATCTTAAGAGACCTTTGAGAAATPAAAGAAATA 155
Oy      121  GCCAGCTGTGAGATGTGTGCTTAAAGCAATCATCACTTCTGTTTATGTTAAAGCCAG 180
Db      156  GCCAGCTGTGAGATGTGTGCTTAAAGCAATCATCACTTCTGTTTATGTTAAAGCCAG 215
Oy      181  AACAGATCTATGAGCAGATTGGCATTCTGGTGTGATCTGTGAGACCAGACTGAGTGC 240
Db      216  AACAGATCTATGAGCAGATTGGCATTCTGGTGTGATCTGTGAGACCAGACTGAGTGC 275
Oy      241  TCCAAAGACCTGAAAAAGCCATCCAAATTAATGTAACCAAAACCTGCAGCAAGATGGCTG 300
Db      276  TCCAAAGACCTGAAAAAGCCATCCAAATTAATGTAACCAAAACCTGCAGCAAGATGGCTG 335
Oy      301  GAGAAAGTTCACTTGGAGCCAGTGAAGTAACCCCACTGGGAGAGGGAGAAATCAAGT 360
Db      336  GAGAAAGTTCACTTGGAGCCAGTGAAGTAACCCCACTGGGAGAGGGAGAAATCAAGT 395
Oy      361  GTGATGCTGAGCCAAAGATTCAATAGATAGCCATCTGGGCTTTGGCAGCAGATTGGG 420
Db      396  GTGATGCTGAGCCAAAGATTCAATAGATAGCCATCTGGGCTTTGGCAGCAGATTGGG 455
Oy      421  ACTCTTCAGAAAGCATTTACAGCAGAAATTTCTGTGTGATGATCTTTGATGAATCTGAG 480
Db      456  ACTCTTCAGAAAGCATTTACAGCAGAAATTTCTGTGTGATGATCTTTGATGAATCTGAG 515
Oy      481  AGAAGGCTTCAGAAAGCAAGAGGAGAAATGTTTATTAACCAACTTACATCACTAC 540
Db      516  AGAAGGCTTCAGAAAGCAAGAGGAGAAATGTTTATTAACCAACTTACATCACTAC 575
Oy      541  TCAAGACGCTGCAATACCCAGACGAGGGCGGTGAGCTGCCAAGTGGGGCTTTTG 600
Db      576  TCAAGACGCTGCAATACCCAGACGAGGGCGGTGAGCTGCCAAGTGGGGCTTTTG 635
Oy      601  GCATCTCTCATTTGATCGATCGTGGCTCTCTTCTCATCTACAGTCTCAACAGTATTCAG 660
Db      636  GCATCTCTCATTTGATCGATCGTGGCTCTCTTCTCATCTACAGTCTCAACAGTATTCAG 695
Oy      661  GAATACAGAGTGGCGGCCCAAAATTCACAGGCTGTATTAAGTGGAGAAATGACGAA 720
Db      696  GAATACAGAGTGGCGGCCCAAAATTCACAGGCTGTATTAAGTGGAGAAATGACGAA 755
Oy      721  ATGATGTCAAGAAATGGCTTCTCATGAGATCAAAATTTGATTCAGCTTAAGATGGGGCA 780
Db      756  ATGATGTCAAGAAATGGCTTCTCATGAGATCAAAATTTGATTCAGCTTAAGATGGGGCA 815
Oy      781  AAGACCTAACCCAGATGATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAATAT 840
Db      816  AAGACCTAACCCAGATGATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAATAT 875
Oy      841  CCAAGACAGTTGTACTGTGTCAGTGAACATCTGACAGCTGGAGATGTTGGCAGAGTGGC 900
Db      876  CCAAGACAGTTGTACTGTGTCAGTGAACATCTGACAGCTGGAGATGTTGGCAGAGTGGC 935
Oy      901  ATGATGATGAGCGGTGAGCCTTATATATCATGAGGAACACTCTCACTTATTAAGATCTT 960
Db      936  ATGATGATGAGCGGTGAGCCTTATATATCATGAGGAACACTCTCACTTATTAAGATCTT 995
Oy      961  GG 962
Db      996  GG 997

RESULT 5
BX439467      1050 bp mRNA linear EST 15-MAY-2003
LOCUS      Homo sapiens PLACENTA Homo sapiens cDNA clone CSDB010YJ06
DEFINITION      BX439467

```

5-PRIME, mRNA sequence.  
 BX439467  
 VERSION BX439467.1 GI:30777746  
 EST.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 1050)  
 L1, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished  
 CONTACT: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4663.f For more information about this cluster, see  
 http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DE010DE03QPI&cluster=4663.f. Contact : Feng liang Email : fliang@life.techn.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DE010DE03QPI.  
 Location/Qualifiers  
 1. 1050  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE010Y006"  
 /tissue\_type="PLACENTA"  
 /clone\_id="Homo sapiens PLACENTA"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."  
 BASE COUNT 307 a 233 c 271 g 239 t  
 ORIGIN  
 Query Match 64.0%; Score 905.8; DB 13; Length 1050;  
 Best Local Similarity 99.8%; Pred. No. 5.2e-254;  
 Matches 907; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 ATGAATTCCTTATCTTGGATTTTGGGTTGTCACCTTTATCCCTGCTGGG 60  
 DB 142 ATGAATTCCTTATCTTGGATTTTGGGTTGTCACCTTTATCCCTGCTGGG 201  
 QY 61 AAAGCTATATGCAAGATGTCATCTTAAGAGACTTTTGAAGAAATTAAGAATA 120  
 DB 202 AAAGCTATATGCAAGATGTCATCTTAAGAGACTTTTGAAGAAATTAAGAATA 261  
 QY 121 GCCAGCTGTGAGATGTTGCTAAAGCAATCAACCTAGCTTTTATGTTAAAGCCG 180  
 DB 262 GCCAGCTGTGAGATGTTGCTAAAGCAATCAACCTAGCTTTTATGTTAAAGCCG 321  
 QY 181 AAGCATCTATGAGCAATGGGCACTCTGTTGATTAAGTGAACCAAGCTAGTGGC 240  
 DB 322 AAGCATCTATGAGCAATGGGCACTCTGTTGATTAAGTGAACCAAGCTAGTGGC 381  
 QY 241 TCCAAAGCACTAAGAAAAGCATCAATTAATACCAAACTGACAGCAATGGGCTG 300  
 DB 382 TCCAAAGCACTAAGAAAAGCATCAATTAATACCAAACTGACAGCAATGGGCTG 441  
 QY 301 GAGAAAGTCACTGAGCCAGTGAATACCCCACTGGAGAGGGGAGAAATCAAGCT 360  
 DB 442 GAGAAAGTCACTGAGCCAGTGAATACCCCACTGGAGAGGGGAGAAATCAAGCT 501  
 QY 361 GTGATGCTGAGCCAAATTCATTAAGTACCAATCCCTGGGCTTTGGAGAGCAATGGG 420  
 DB 502 GTGATGCTGAGCCAAATTCATTAAGTACCAATCCCTGGGCTTTGGAGAGCAATGGG 561  
 QY 421 ACTCTCCAGAAAGCATTAAGAGAGTCTGCTGATGAAGCTTTTCATGAATGAG 480

|||||  
 DB 562 ACTCTCCAGAAAGCATTAAGAGAGTCTGCTGATGAAGCTTTTCATGAATGAG 621  
 QY 481 AGAAGGCTCCAGAAAGCAAGGAAGATGTTGTTATTAACCACTTATCAATAC 540  
 DB 622 AGAAGGCTCCAGAAAGCAAGGAAGATGTTGTTATTAACCACTTATCAATAC 681  
 QY 541 TCAAGAGCTGCAATACCAAGCAAGGAGGCGGTGAGAGCTGCCAAGTGGGCTTTG 600  
 DB 682 TCAAGAGCTGCAATACCAAGCAAGGAGGCGGTGAGAGCTGCCAAGTGGGCTTTG 741  
 QY 601 GCATCTCTCATTCGATCCGTTGCGCTCTCTTCAATCAAGTCTCAACAGTATTG 660  
 DB 742 GCATCTCTCATTCGATCCGTTGCGCTCTCTTCAATCAAGTCTCAACAGTATTG 801  
 QY 661 GAATACAGATGAGCGGCCCAAAATTCACCAAGCTGTTATTAAGTGAAGATGCA 720  
 DB 802 GAATACAGATGAGCGGCCCAAAATTCACCAAGCTGTTATTAAGTGAAGATGCA 861  
 QY 721 ATGATGTCAGAAATGCTTCTCATGGGATCAAAATGTCATTGCAATGAAGTGGGCA 780  
 DB 862 ATGATGTCAGAAATGCTTCTCATGGGATCAAAATGTCATTGCAATGAAGTGGGCA 921  
 QY 781 AAGACTTACCCAGATGATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCA 840  
 DB 922 AAGACTTACCCAGATGATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCA 981  
 QY 841 CCAGAACAGTTGTACTGCTGTCAGTGAACATCTGACAGCTGGATTTGGGAGGTC 900  
 DB 982 CCAGAACAGTTGTACTGCTGTCAGTGAACATCTGACAGCTGGATTTGGGAGGTC 1041  
 QY 901 ATGATGAT 909  
 DB 1042 ATGATGAT 1050  
 RESULT 6  
 LOCUS BX355940  
 DEFINITION BX355940 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS0D1006Y115 5-PRIME, mRNA sequence.  
 ACCESSION BX355940  
 VERSION BX355940.1 GI:30384019  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 1038)  
 L1, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished  
 CONTACT: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4663.f For more information about this cluster, see  
 http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1006AB08QPI&cluster=4663.f. Contact : Feng liang Email : fliang@life.techn.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0D1006AB08QPI.  
 Location/Qualifiers  
 1. 1038  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1006Y115"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_id="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)"

primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 304 a 237 c 262 g 233 t 2 others

Query Match 59.6%; Score 843.8; DB 13; Length 1038;  
Best Local Similarity 99.7%; Pred. No. 8.1e-236;  
Matches 856; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATGAATTCCTTATCTTGGCAATTTTGGGTGTTCACTTTATCCCTGCTCTGGG 60  
DB 180 ATGAATTCCTTATCTTGGCAATTTTGGGTGTTCACTTTATCCCTGCTCTGGG 239  
QY 61 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAATA 120  
DB 240 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAATA 299  
QY 121 GCCAGCTGTGAGATGTTGCTAAGCAATCATCACTAGCTGTTTATGTAAGCCGAG 180  
DB 300 GCCAGCTGTGAGATGTTGCTAAGCAATCATCACTAGCTGTTTATGTAAGCCGAG 359  
QY 181 AACAGATCTTATGAGCGATGGCACTTCTGTGTTGACTGTGACCCAGACTAGTGGC 240  
DB 360 AACAGATCTTATGAGCGATGGCACTTCTGTGTTGACTGTGACCCAGACTAGTGGC 419  
QY 241 TCCAGAACTTATGAGAAAGCCATTCGAATTTATGTAACCAAACTTCGACAGATGGGCTG 300  
DB 420 TCCAGAACTTATGAGAAAGCCATTCGAATTTATGTAACCAAACTTCGACAGATGGGCTG 479  
QY 301 GAGAAAGTTTCACTGAGAGCCAGTGAATATACCCCACTGGGAGAGGGAGAAATCAGCT 360  
DB 480 GAGAAAGTTTCACTGAGAGCCAGTGAATATACCCCACTGGGAGAGGGAGAAATCAGCT 539  
QY 361 GTGATCTGAGAGCCAGAAATTCATTAAGATAGCATCTTGGGTCTTGGAGAGCATTTGGG 420  
DB 540 GTGATCTGAGAGCCAGAAATTCATTAAGATAGCATCTTGGGTCTTGGAGAGCATTTGGG 599  
QY 421 ACTCTCCAGAAAGCATTAAGCAAGATTTCTGTGTTGACTCTTTGATGAATTTGGAG 480  
DB 600 ACTCTCCAGAAAGCATTAAGCAAGATTTCTGTGTTGACTCTTTGATGAATTTGGAG 659  
QY 481 AGAAGGGCTCTGAGAGCAAGAGGAGATTTGTTTAAACCACTTACATCACTAC 540  
DB 660 AGAAGGGCTCTGAGAGCAAGAGGAGATTTGTTTAAACCACTTACATCACTAC 719  
QY 541 TCAAGGACGGTGCATTAATCCGAAACGAGGGGGGCTGTGAAAGCTGCCAAGGTGGGCTTTG 600  
DB 720 TCAAGGACGGTGCATTAATCCGAAACGAGGGGGGCTGTGAAAGCTGCCAAGGTGGGCTTTG 779  
QY 601 GCATCTCTCATTCGATCCGTTGGCTCTTCTTCATCTTCAAGTCTTCACACAGTATTAG 660  
DB 780 GCATCTCTCATTCGATCCGTTGGCTCTTCTTCATCTTCAAGTCTTCACACAGTATTAG 839  
QY 661 GAATACCAAGATGGGCTGCCCAAAATTCACACAGCTGTATTAACGTTGAAGATCAGAA 720  
DB 840 GAATACCAAGATGGGCTGCCCAAAATTCACACAGCTGTATTAACGTTGAAGATCAGAA 899  
QY 721 ATGATGTCAAGATGGCTTCTCATGGGATCAAAATTTGCTACCTTAAGAT-GGGGGC 779  
DB 900 ATGATGTCAAGATGGCTTCTCATGGGATCAAAATTTGCTACCTTAAGATGGGGGGC 959  
QY 780 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAATA 120  
DB 960 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAATA 1019  
QY 840 TCCAGAACTTATGAGCGATGGCACTTCTGTGTTGACTGTGACCCAGACTAGTGGC 240  
DB 1020 TCCAGAACTTATGAGCGATGGCACTTCTGTGTTGACTGTGACCCAGACTAGTGGC 300

RESULT 7  
B0878966

LOCUS B0878966 909 bp mRNA linear EST 16-AUG-2002  
DEFINITION AGENCOURT 8183107 Lupski dorsal root ganglion Homo sapiens cDNA  
clone IMAGE:6184129 5', mRNA sequence.  
ACCESSION B0878966  
VERSION B0878966.1 GI:22270974  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 909)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Dr. James R. Lupski  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LLM13572 row: n column: 02  
High quality sequence stop: 696.  
Location/Qualifiers  
1..909  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6184129"  
/sex="male"  
/tissue\_type="dorsal root ganglia"  
/dev\_stage="adult, 36 yr"  
/lab\_host="DH10B"  
/clone\_id="Lupski\_dorsal\_root\_ganglion"  
/note="Vector: pCMV-SPORT6 (Life Technologies); Site\_1:  
NotI; Site\_2: SalI; cDNA made by oligo-dT priming.  
Directionally cloned using the following adaptors:  
5'-TCGACCCAGGCGTCCG-3' and  
5'-GACTAGTTCGATGCGAGCGGCGCCCT(15)-3'. Size selected >  
1 kb for average insert length 1.7 kb. This is a primary  
library, non-amplified. Library constructed by Life  
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
College of Medicine) and is available through Life  
Technologies."

BASE COUNT 273 a 202 c 226 g 207 t 1 others

Query Match 56.3%; Score 796.8; DB 13; Length 909;  
Best Local Similarity 98.8%; Pred. No. 4.8e-222;  
Matches 823; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 1 ATGAATTCCTTATCTTGGCAATTTTGGGTGTTCACTTTATCCCTGCTCTGGG 60  
DB 64 ATGAATTCCTTATCTTGGCAATTTTGGGTGTTCACTTTATCCCTGCTCTGGG 123  
QY 61 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAATA 120  
DB 124 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAATA 183  
QY 121 GCCAGCTGTGAGATGTTGCTAAGCAATCATCACTAGCTGTTTATGTAAGCCGAG 180  
DB 184 GCCAGCTGTGAGATGTTGCTAAGCAATCATCACTAGCTGTTTATGTAAGCCGAG 243  
QY 181 AACAGATCTTATGAGCGATGGCACTTCTGTGTTGACTGTGACCCAGACTAGTGGC 240  
DB 244 AACAGATCTTATGAGCGATGGCACTTCTGTGTTGACTGTGACCCAGACTAGTGGC 303  
QY 241 TCCAGAACTTATGAGCGATGGCACTTCTGTGTTGACTGTGACCCAGACTAGTGGC 300  
DB 304 TCCAGAACTTATGAGCGATGGCACTTCTGTGTTGACTGTGACCCAGACTAGTGGC 363

```

QY 301 GAGAAATTCACTGAGCCAGTGAATATCCCACTGGAGAGAGAGAAATCAAGCT 360
Db 364 GAGAAATTCACTGAGCCAGTGAATATCCCACTGGAGAGAGAGAAATCAAGCT 423
QY 361 GTGATGTGAGCCGAATATTCATTAATAGTCCATCTCGGTCTTGGCAGAGATTTGG 420
Db 424 GTGATGTGAGCCGAATATTCATTAATAGTCCATCTCGGTCTTGGCAGAGATTTGG 483
QY 421 ACTCTCTCAGAAAGCATTTACAGCAGAAATTTCTGTGTGACCTTTTCATGATCTGAG 480
Db 484 ACTCTCTCAGAAAGCATTTACAGCAGAAATTTCTGTGTGACCTTTTCATGATCTGAG 543
QY 481 AGAAGGCTCTCAGAAAGCATTTACAGCAGAAATTTCTGTGTGACCTTTTCATGATCTGAG 540
Db 544 AGAAGGCTCTCAGAAAGCATTTACAGCAGAAATTTCTGTGTGACCTTTTCATGATCTGAG 603
QY 541 TCAAGAGCGGTGCAATACCAAGCCAGAGGGGGGGTGGAGCTGCCAAGGTGGGGCTTTG 600
Db 604 TCAAGAGCGGTGCAATACCAAGCCAGAGGGGGGGTGGAGCTGCCAAGGTGGGGCTTTG 663
QY 601 GCATCTCTCATTCGATCCGTGGCTCTCTTCATCTACAGTCTCTCAGACAGGATTCAG 660
Db 664 GCATCTCTCATTCGATCCGTGGCTCTCTTCATCTACAGTCTCTCAGACAGGATTCAG 723
QY 661 GAATACCAAGATGGCGTGGCCCAAAATTCACAAGCTGTATTAAGTGAAGATGCAGAA 720
Db 724 GAATACCAAGATGGCGTGGCCCAAAATTCACAAGCTGTATTAAGTGAAGATGCAGAA 783
QY 721 ATGATGTCAAGATGGCTTTCTCATGGATCAAAATTTGTCAATCACTTAAGAT- 779
Db 784 ATGATGTCAAGATGGCTTTCTCATGGATCAAAATTTGTCAATCACTTAAGATGGGGGCG 843
QY 780 AAAGACCTACCCAGATCTGATCTTCAACACT-GTAGCAGAGATCACTGGG 831
Db 844 AAAGACCTACCCAGATCTGATCTTCAACACTGTAGCAGAGATCACTGGG 896

```

```

RESULT 8
BX360507 1126 bp mRNA linear EST 05-MAY-2003
LOCUS BX360507
DEFINITION clone CSOD1071A19 5-PRIME, mRNA sequence.
ACCESSION BX360507
VERSION BX360507.1 GI:30374439
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1126)
LJ.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSOD1071A19001&cluster=4663.f. Contact:
Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CSOD1071A19001.
location/Qualifiers
1. 1126
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1071A19"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"

```

## FEATURES

source

```

1. 1126
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1071A19"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"

```

```

/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 311 a 249 c 298 g 246 t 22 others
ORIGIN

```

```

Query Match 54.7%; Score 774.4; DB 13; Length 1126;
Best Local Similarity 96.6%; Pred. No. 2e-215;
Matches 833; Conservative 9; Mismatches 15; Indels 5; Gaps 5;

```

```

QY 1 ATGAATTCCTTATCTTGGCATTCTTGGTGTGTTACCTTTATCCTGTGCTGGG 60
Db 210 ATGAATTCCTTATCTTGGCATTCTTGGTGTGTTACCTTTATCCTGTGCTGGG 269
QY 61 AAGCTATATGCAAGAAATGGCATCTCTAAGAGAACTTTGAAGAAATTAAGAAATA 120
Db 270 AAGCTATATGCAAGAAATGGCATCTCTAAGAGAACTTTGAAGAAATTAAGAAATA 329
QY 121 GCCAGCTGTGAGATGTTGCTTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 180
Db 330 GCCAGCTGTGAGATGTTGCTTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 389
QY 181 AACAGATCTATAGCGATTTGGCACTTGTGTTGATACCTGTTGACCCAGACTGAGTGGC 240
Db 390 AACAGATCTATAGCGATTTGGCACTTGTGTTGATACCTGTTGACCCAGACTGAGTGGC 449
QY 241 TCCAGAACTTAAGAAAGGCAATCCAAATTTATGATCCAAACCTGCAGCAAGTGGCTG 300
Db 450 TCCAGAACTTAAGAAAGGCAATCCAAATTTATGATCCAAACCTGCAGCAAGTGGCTG 509
QY 301 GAGAAATTCACCTGGAGCCAGTGAATATCCCACTGGGAGAGGGAGAAATCAAGCT 360
Db 510 GAGAAATTCACCTGGAGCCAGTGAATATCCCACTGGGAGAGGGAGAAATCAAGCT 569
QY 361 GTGATGTGAGCCGAATATTCATTAAGATAGCCATCTGGGCTTGGCAGCAGATTTGGG 420
Db 570 GTGATGTGAGCCGAATATTCATTAAGATAGCCATCTGGGCTTGGCAGCAGATTTGGG 629
QY 421 ACTCTCTCAGAAAGCATTTACAGCAGAAATTTCTGTGTGACCTTTTCATGATCTGAG 480
Db 630 ACTCTCTCAGAAAGCATTTACAGCAGAAATTTCTGTGTGACCTTTTCATGATCTGAG 689
QY 481 AGAAGGCTCTCAGAAAGCATTTACAGCAGAAATTTCTGTGTGACCTTTTCATGATCTGAG 540
Db 690 AGAAGGCTCTCAGAAAGCATTTACAGCAGAAATTTCTGTGTGACCTTTTCATGATCTGAG 749
QY 541 TCAAGAGCGGTGCAATACCAAGCCAGAGGGGGGGTGGAGCTGCCAAGGTGGGGCTTTG 600
Db 750 TCAAGAGCGGTGCAATACCAAGCCAGAGGGGGGGTGGAGCTGCCAAGGTGGGGCTTTG 809
QY 601 GCATCTCTCATTCGATCCGTGGCTCTCTTCATCTACAGTCTCTCAGACAGGATTCAG 660
Db 810 GCATCTCTCATTCGATCCGTGGCTCTCTTCATCTACAGTCTCTCAGACAGGATTCAG 869
QY 661 GAATACCAAGATGGCGTGGCCCAAAATTCACAAGCCGTGTTATAGCACTTACATCAACTAC 720
Db 870 GAATACCAAGATGGCGTGGCCCAAAATTCACAAGCCGTGTTATAGCACTTACATCAACTAC 929
QY 720 AATGATGTCAAGATGGCTTTCTCA-TGGATCAAAATTTGTCAATCACTTAAGATGGGGG 778
Db 930 AATGATGTCAAGATGGCTTTCTCA-TGGATCAAAATTTGTCAATCACTTAAGATGGGGG 989
QY 779 CAAGACTTACCAAGATGATGATTTCTTCAACACTGTAGAGAGATCACTGGAGCAAT 838
Db 990 S-AARAMCTACCAAGATGATGATTTCTTCAACACTGTAGAGAGATCACTGGAGCAAMA 1046
QY 839 ATCCAGAAAGCTGTACTGCT 860
Db 1047 TCARAAAGGTACTGTAAGT 1068

```

RESULT 9



BX335995  
LOCUS BX335995 919 bp mRNA linear EST 02-MAY-2003  
DEFINITION BX335995 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0D1022YU16 5-PRIME, mRNA sequence.  
ACCESSION BX335995  
VERSION BX335995.1 GI:30339459  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 919)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91060 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4663.f For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0D1022DE08QPlac1cluster=4663.f. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0D1022DE08QPl.  
Location/Qualifiers  
1. 919  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1022YU16"  
/rissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 273 a 206 c 226 g 208 t  
ORIGIN  
Query Match 54.1%; Score 766.2; DB 13; Length 919;  
Best Local Similarity 98.5%; Pred. No. 4.6e-213;  
Matches 779; Conservative 4; Mismatches 7; Indels 1; Gaps 1;  
QY 1 ATGAATTCCTTATCTTGCATTTTTCGCTGCTTCACTTTTATCCCTGCTCTGGG 60  
DB 126 ATGAATTCCTTATCTTGCATTTTTCGCTGCTTCACTTTTATCCCTGCTCTGGG 185  
QY 61 AAAGCTATATGCAAGATGCGATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATTA 120  
DB 186 AAAGCTATATGCAAGATGCGATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATTA 245  
QY 121 GCCAGCTGTGAGATGTTGTTAAACAATCACTAGCTGTTTATGTTAAAGCCAG 180  
DB 246 GCCAGCTGTGAGATGTTGTTAAACAATCACTAGCTGTTTATGTTAAAGCCAG 305  
QY 181 AACAGATCTTATAGCGATTGGCACTTCTGTGTTAGTCTTTGACCCAGACTAGTGGC 240  
DB 306 AACAGATCTTATAGCGATTGGCACTTCTGTGTTAGTCTTTGACCCAGACTAGTGGC 365  
QY 241 TCCAGAACCTTATAGCGATTGGCACTTCTGTGTTAGTCTTTGACCCAGACTAGTGGC 300  
DB 366 TCCAGAACCTTATAGCGATTGGCACTTCTGTGTTAGTCTTTGACCCAGACTAGTGGC 425  
QY 301 GAGAAAGTTCACTGTGAGAGCCAGTGAATAACCCACTGGAGAGGAGGAGAAATCACT 360  
DB 426 GAGAAAGTTCACTGTGAGAGCCAGTGAATAACCCACTGGAGAGGAGGAGAAATCACT 485  
QY 361 GTGATGCTGAGAGCCAGATTCATTAAGATAGCCATCTGGGCTTTGGAGCAGCATTTGG 420  
DB 486 GTGATGCTGAGAGCCAGATTCATTAAGATAGCCATCTGGGCTTTGGAGCAGCATTTGG 545

QY 421 ACTCTCCAGAAAGCATTTACAGCAGAGATTCTGGTGTGACCTCTTTCAGTAAGTCAG 480  
DB 546 ACTCTCCAGAAAGCATTTACAGCAGAGATTCTGGTGTGACCTCTTTCAGTAAGTCAG 605  
QY 481 AGAAGGCTTCAGAAAGCAGAGAGAGATTGTTTATTAACAACCTTATCATCACTAC 540  
DB 606 AGAAGGCTTCAGAAAGCAGAGAGAGATTGTTTATTAACAACCTTATCATCACTAC 665  
QY 541 TCAAGACGCTGCAATCCGACGAGGCGGTGGAACCTGCAAGGTGGGGCTTTG 600  
DB 666 TCAAGACGCTGCAATCCGACGAGGCGGTGGAACCTGCAAGGTGGGGCTTTG 725  
QY 601 GCATCTCATTCATTCGATCCGTGCGCTCTCTTCATCTACAGTCCCTCACAGGATTCAG 660  
DB 726 GCATCTCATTCATTCGATCCGTGCGCTCTCTTCATCTACAGTCCCTCACAGGATTCAG 785  
QY 661 GAATACAGAGATGCGTGGCCCAAAATTCACAACGCTGTATTCGTTGAGATGCGAA 720  
DB 786 GAATACAGAGATGCGTGGCCCAAGATTCCAACGCTGTATTCGTTGAGATGCGAA 845  
QY 721 ATGATGTCAGAAATGCGTTCTCAT -GGATCAAAATGTCATTCAGTAAAGATGGGGC 779  
DB 846 ATGATGTCAGAAATGCGTTCTCATGGGATCAAAATGTCATTCAGTAAAGATGGGGC 905  
QY 780 AAAGACCTACC 790  
DB 906 AAAGACCTACC 916  
RESULT 10  
BU146905 866 bp mRNA linear EST 03-SEP-2002  
LOCUS BU146905  
DEFINITION AGENCOURT 7978204 Lupski dorsal\_root\_ganglion Homo sapiens cDNA  
clone IMAGB:6185233 5', mRNA sequence.  
ACCESSION BU146905  
VERSION BU146905.1 GI:22660437  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 866)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Dr. James R. Lupski  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
plate: LPLM13575 row: 1 column: 02  
High quality sequence stop: 731.  
Location/Qualifiers  
1. 866  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGB:6185233"  
/sex="male"  
/rissue\_type="dorsal root ganglion"  
/dev\_stage="adult, 36 yr"  
/lab\_host="DH10B"  
/clone\_lib="Lupski dorsal root ganglion"  
/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:  
NotI; Site 2: SalI; cDNA made by oligo-dT priming.  
Directionally cloned using the following adaptors:  
5'-TCGACCCAGCGCTCCG-3' and  
5'-GACTAGTTTATGATCGCAGGCGCCCTT (15)-3'. Size selected >



1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies.

BASE COUNT 257 a 182 c 233 g 191 t 3 others

Query Match 52.9%; Score 748.6; DB 13; Length 866;  
Best Local Similarity 99.1%; Pred. No. 6.6e-208;  
Matches 773; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

65 CTATATGCAAGAAATGGCATCTCTTAAGAGACTTT-TGAGAAATTAAGAAATAGGCC 123  
1 CTATATGCAAGAAATGGCATCTCTTAAGAGACTTTTGAAGAAATTAAGAAATAGGCC 60  
124 AGCTGTGAGAGTGTGCTTAAGCAATCAACCTAGCTGTTATGTAAAGCCAGAAC 183  
61 AGCTGTGAGAGTGTGCTTAAGCAATCAACCTAGCTGTTATGTAAAGCCAGAAC 120  
184 AGATCTATGAGAGTGTGCTTCTGTGTGATCTGTGTGAACCCAGCTGAGCTCC 243  
121 AGATCTATGAGAGTGTGCTTCTGTGTGATCTGTGTGAACCCAGCTGAGCTCC 180  
244 AAGAACTAGAAAGAAAGCCATCCAAATTAATGTAACAAACCTGACAGCAAGATGGCTGGAG 303  
181 AAGAACTAGAAAGAAAGCCATCCAAATTAATGTAACAAACCTGACAGCAAGATGGCTGGAG 240  
304 AAGATTCACCTGAGAGCCAGTGAATATCCCACTGGAGAGAGGAGAAATCAGCTGTG 363  
241 AAGATTCACCTGAGAGCCAGTGAATATCCCACTGGAGAGAGGAGAAATCAGCTGTG 300  
364 ATGTGTGAGAGCAAAATTCATTAAGATAGCCATCCCTGGCTCTTGTGAGAGAGATGGGACT 423  
301 ATGTGTGAGAGCAAAATTCATTAAGATAGCCATCCCTGGCTCTTGTGAGAGAGATGGGACT 360  
424 CCTCAGAGAGGCAATTAACAGAGATTTGTGTGTGATCCTCTTTCAGATGAATCTGAGAGA 483  
361 CCTCAGAGAGGCAATTAACAGAGATTTGTGTGTGATCCTCTTTCAGATGAATCTGAGAGA 420  
484 AGGAGCTCAGAGAGCAAGAGAGAGATTTGTGTGTGATCCTCTTTCAGATGAATCTGAGAGA 543  
421 AGGAGCTCAGAGAGCAAGAGAGAGATTTGTGTGTGATCCTCTTTCAGATGAATCTGAGAGA 480  
544 AGAGAGGTCGAATACCGAAGCGAGGGGCGGTGAGAGCTGCCAGATGGGGCTTTGGCA 603  
481 AGAGAGGTCGAATACCGAAGCGAGGGGCGGTGAGAGCTGCCAGATGGGGCTTTGGCA 540  
604 TCTCTCATTCGATCCGTGGCTCCTCTTCATCTAAGATCTGCTCAACAGATTAATCAGGAA 663  
541 TCTCTCATTCGATCCGTGGCTCCTCTTCATCTAAGATCTGCTCAACAGATTAATCAGGAA 600  
664 TACAGAGATGGCGTGGCCCAAAATTCACAGAGCTGTATTAAGGTGAGAGATGACAGAAATG 723  
601 TACAGAGATGGCGTGGCCCAAAATTCACAGAGCTGTATTAAGGTGAGAGATGACAGAAATG 660  
724 ATGTCAAGATGGCTTCTCATGAGGATCAAAATTTGATTCAGCTTAAGATGGGGCAGAAAG 783  
661 ATGTCAAGATGGCTTCTCATGAGGATCAAAATTTGATTCAGCTTAAGATGGGGCAGAAAG 720  
784 ACTTACCAAGATTAATCTGCTTCAACACT-CTACAGAGATCACTGGGAGCAAAATATCC 842  
721 ACTTACCAAGATTAATCTGCTTCAACACTGCTGAGAGATCACTGGGAGCAAAATATTC 780

RESULT 11  
BX416896 1012 bp mRNA linear EST 13-MAY-2003  
LOCUS BX416896 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE001YF04  
DEFINITION 5-PRIME, mRNA sequence.  
ACCESSION BX416896  
VERSION BX416896.1 GI:30650313  
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1012)

Li, W.B., Gruber, C., Jesse, J. and Polayes, D. Full-length cDNA libraries and normalization

Unpublished

JOURNAL

Contact: Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 4663.f For  
more information about this cluster, see

http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DE001DC02QPIcluster=4663.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paradise Avenue Genoscope sequence ID : CS0DE001DC02QPI.  
Location/Qualifiers

FEATURES  
Source  
1..1012  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE001YF04"  
/issue\_type="PLACENTA"  
/clone\_id="Homo sapiens PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo (dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

BASE COUNT 281 a 240 c 275 g 215 t 1 others  
ORIGIN  
Query Match 52.5%; Score 742.8; DB 13; Length 1012;  
Best Local Similarity 99.6%; Pred. No. 3.5e-206;  
Matches 755; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
1 ATGAAATTCCTATCTTGGCAATTTTGGTGGTTCACCTTTATCCCTGTGCTGGG 60  
255 ATGAAATTCCTATCTTGGCAATTTTGGTGGTTCACCTTTATCCCTGTGCTGGG 314  
61 AAAGCTATGCAAGATGGCATCTTAAGAGACTTTTGAAGAAATTAAGAAATTA 120  
315 AAAGCTATGCAAGATGGCATCTTAAGAGACTTTTGAAGAAATTAAGAAATTA 374  
121 GCCAGCTGTGAGATGTGCTTAAGCAATCAACCTAGCTGTTATGTAAAGCCGAG 180  
375 GCCAGCTGTGAGATGTGCTTAAGCAATCAACCTAGCTGTTATGTAAAGCCGAG 434  
181 AACAGATCTATGAGAGATTTGGCACTTCTGGTGTATCTTTGAGCCAGATGAGGC 240  
435 AACAGATCTATGAGAGATTTGGCACTTCTGGTGTATCTTTGAGCCAGATGAGGC 494  
241 TCAGAGAACTTAAGAAAGCCATCCAAATTAATGTAACAAACCTGACAGAAATGGGCTG 300  
495 TCAGAGAACTTAAGAAAGCCATCCAAATTAATGTAACAAACCTGACAGAAATGGGCTG 554  
301 GAGAAAGTTTCACTGAGAGCCAGTGAATACCCCACTGGGAGAGAGGAGAAATCAGCT 360  
555 GAGAAAGTTTCACTGAGAGCCAGTGAATACCCCACTGGGAGAGAGGAGAAATCAGCT 614  
361 GTGATGTGAGAGCCAGAAATTCATTAAGATAGCCATCTGAGTCTTTGGCAGCAGATTGGG 420  
615 GTGATGTGAGAGCCAGAAATTCATTAAGATAGCCATCTGAGTCTTTGGCAGCAGATTGGG 674  
421 ACTCTTCCAGAGGCAATTAACAGCAAGATTTGTGTGTGATCCTCTTTGATGAATCTGAG 480  
675 ACTCTTCCAGAGGCAATTAACAGCAAGATTTGTGTGTGATCCTCTTTGATGAATCTGAG 734  
481 AGAAGGGCTCAGAGAGCAAGAGAGATTTGTGTGTATTAACCAACTTATACATCACTAC 540

Db	735	AGAAAGGCGCTCAGAGCAAGAGGAAGATTGTTATTATACCAACCTTATCATCTCAACTAC	794
Qy	541	TCAAGGACGGTGTGCATTATCCGAAACGAGGGGGGGGTGGAGACGTGCCAAGG-TGGGGGCTTT	599
Db	795	TCAGAGACGGTGTGCATTATCCGAAACGAGGGGGGGGGTGAAGCTGCCAAGGTTGGGGCTTT	854
Qy	600	GGCATCTCTCATTTGCATCCGTGGGCTCTCTTCTTCATCTTACAGTCTCTCACACAGGTAATTC	659
Db	855	GGCATCTCTCATTTGCATCCGTGGGCTCTCTTCTTCATCTTACAGTCTCTCACACAGGTAATTC	914
Qy	660	GGAAATACCAAGATGGCGTCCCAAATTTCCACAGCGCTGTATTAACGTGGGAAGATGCAGA	719
Db	915	GGAAATACCAAGATGGCGTCCCAAATTTCCACAGCGCTGTATTAACGTGGGAAGATGCAGA	974
Qy	720	AATGATGTCAAGAAATGGCTTCTTCATAGGAGTCAAATTTG	757
Db	975	AATGATGTCAAGAAATGGCTTCTTCATAGGAGTCAAATTTG	1012

RESULT 12	
LOCUS	6G751497
DEFINITION	6G751497 902 bp mRNA linear EST 15-MAY-2001 602730076E1 NIH_MGC_43 Homo sapiens CDNA clone IMAGE:4873770 5',

ACCESSION	BG751497
VERSION	BG751497.1
	GI:14062150

SOURCE ORGANISM	Homo sapiens (human)
	Homo sapiens

REFERENCE	
AUTHORS	Eukariyola; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 902)
JOURNAL	NIH-MGC <a href="http://mgc.ncbi.nlm.nih.gov/">http://mgc.ncbi.nlm.nih.gov/</a> .
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished
	Contact: Robert Strausberg, Ph.D.

Email: [cgaps-i@mail.nih.gov](mailto:cgaps-i@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LCM1752 row: 9 column: 19  
High quality sequence stop: 817.

## FEATURES

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4873770"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_1b="NIH_MGC_43"
/notes="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
Note: This is a NIH_MGC Library. |"
BASE COUNT
246 a 200 c 251 g 204 t 1 others

```

```

Query Match      52.3%;   Score 740;   DB 10;   Length 902;
Best Local Similarity 95.7%;   Pred. No. 2.2e-205;
Matches 84%;   Conservative 26;   Indels 12;   Gaps 8;

QY      331  CCCCACTGGGAGAGGGGAGAGAAGTCACTGTGATGATCGAGAGCCAAAGATTATTAAGATA 390
|||||

```

Db	2	CCCCACTGGGAGAGGGGAGAAAGTA	CAAGCTGATGCTGGAGGCCAAGAATTCA	TAAAGTA	61
OY		391	GCCATCCTGGGCTTTGGCAGCAGCATTTGGGACTCTCTCCAGAAAGCATTTACGCGAAAGTT		450
Db		62	GCAATCCCTGGGCTTTGGCAGCAGCATTTGGGACTCTCTCCAGAAAGCATTTACGCGAAAGTT		121
OY		451	CTGGTGGGACCTCTTTTCGATGAATCTCAGAGAGGGCTCAGAAAGCAAGAGGGAAGTT		510
Db		122	CTGGTGGGACCTCTTTTCGATGAATCTCAGAGAGAGGGCTCAGAAAGCAAGAGGGAAG-T		180
OY		511	GTTGTTTAAACCAACCTTACATCACTA	CTCAAGACGCTGCAATACCGAACCGAGGG	570
Db		181	GTTGTTTAAACCAACCTTACATCACTA	CTCAAGACGCTGCAATACCGAACCGAGGG	240
OY		571	GGGGTGAAGGTGGCAAGGTGGGGGCTTTGGCATCTCTCA	TTGGATCCGTTGGGCTCTCTTC	630
Db		241	GGGGTGAAGGTGGCAAGGTGGGGGCTTTGGCATCTCTCA	TTGGATCCGTTGGGCTCTCTTC	300
OY		631	TCCATCTACAGTCCCTCACACAGATATT	CAGAAATACAGATGGCGTGGCCAAATTTCA	690
Db		301	TCCATCTACAGTCCCTCACACAGATATT	CAGAAATACAGATGGCGTGGCCAAATTTCA	360
OY		691	ACAGCCTGTATTACGTTGGAAGATGCGAAATGATGTC	AAAGATGGCTTCTCATGGGATC	750
Db		361	ACAGCCTGTATTACGTTGGAAGATGCGAAATGATGTC	AAAGATGGCTTCTCATGGGATC	420
OY		751	AAATTTGTCATTCAAGCTAAAGATGGGGGCAAAAGACTAC	CCAGATCTGATTTCTCTCAAC	810
Db		421	AAATTTGTCATTCAAGCTAAAGATGGGGGCAAAAGACTAC	CCAGATCTGATTTCTCTCAAC	480
OY		811	ACTGTAGCAGAGATCACTGGGAGGCAAAATATC	CAGAACAGTTGTACTGGTCAGTGCAT	870
Db		481	ACTGTAGCAGAGATCACTGGGAGGCAAAATATC	CAGAACAGTTGTACTGGTCAGTGCAT	540
OY		871	CTGGAACAGCTGGGATGTTGGGCAAGGGTGCATGG	-ATGATGGCGGTGAAGCCTTATATC	929
Db		541	CTGGAACAGCTGGGATGTTGGGCAAGGGTGCATGG	-ATGATGGCGGTGAAGCCTTATATC	600
OY		930	ATGGGAAGCACTCTCACTTATTTAAAGATCTTTGGGCTCGT	CCAAAGAGACTCTCGGACT	989
Db		601	ATGGGAAGCACTCTCACTTATTTAAAGATCTTTGGGCTCGT	CCAAAGAGACTCTCGGACT	660
OY		990	GGTGCTCTGGACT-GCGAAGAACAGGTGAGTTGTGCTT	CCAGTATTTATCAATTAC	1048
Db		661	GGTGCTCTGGACTGGCGAAGAACAGGTGAGTTGTGCTT	CCAGTATTTATCAATTAC	720
OY		1049	ACAAGGTAAATATTTCACACTA--CAGTCTGGGTATGAGATC	-TGACGCGAAGAACTTCT	1105
Db		721	ACAAGGTAAATATTTCACACTAACA	GTCGTGGGTATGAGTCTTGACGCGAAGAACTTCT	780
OY		1106	T--ACCCACTGGGCTGCAATTCC	-TGGCAGTGAAGAGCCAGAGGCTATCATGAGAG	1161
Db		781	TTAACCCACTTTGGGCTCAATTCTCACTTGGCAGTGA	AAAAAGGCCAGGCTATCATGAGAG	840
OY		1162	GTTATGAGCCTCTGGACCCCTCAATATATCACTCAAGT	CTCTGGAGC	1206
Db		841	GTTATGAGCCTCTGATGAAGCCC--TCATATCACTCAAGT	CTCGAACC	883

RESULT 13					
LOCUS	BX460463				
DEFINITION	BX460463	998 bp	mRNA	linear	EST 22-MAY-2003
ACCESSION	BX460463	Homo sapiens	PETAL BRAIN	Homo sapiens	cDNA clone
VERSION	CS00F014YN05	5-PRIME	mRNA sequence.		
KEYWORDS	BX460463				
SOURCE	BX460463.1	GI:31029218	EST.		
ORGANISM	Homo sapiens (human)				
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Tesse, J. and Polayes, D.				

TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Genoscope -- Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4663.f For more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CSODF014CG03QPLcluster=4663.f. Contact :  
Peng Liang Email : filiang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CSODF014CG03QPL.

FEATURES  
SOURCE  
1. 998  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODF014YN05"  
/tissue\_type="FETAL BRAIN"  
/dev\_stage="fetal"  
/clone\_lib="Homo sapiens FETAL BRAIN"  
/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

BASE COUNT  
280 a 232 c 233 g 223 t 30 others

Query Match 52.0%; Score 736.6; DB 13; Length 998;  
Best Local Similarity 97.4%; Pred. No. 2.3e-204; Indels 0; Gaps 0;  
Matches 736; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 1 ATGAAATCTCTTATCTTGGATTTTGGGATGTTCACTTTATCCCTGCTGGG 60  
DB 239 ATGAAATCTCTTATCTTGGATTTTGGGATGTTCACTTTATCCCTGCTGGG 298

QY 61 AAAGCTATTCAGAAATGCGATCTCTAAGAGCATTTTGAAGAATAAAGAATA 120  
DB 299 AAAGCTATTCAGAAATGCGATCTCTAAGAGCATTTTGAAGAATAAAGAATA 358

QY 121 GCCAGCTGTGAGATGTTGTTAAGCAATCATCACTAGCTTTATGTTAAGCCGAG 180  
DB 359 GCCAGCTGTGAGATGTTGTTAAGCAATCATCACTAGCTTTATGTTAAGCCGAG 418

QY 181 AACAGATCTATGAGCGATTTGGCACTTCTGTTGATCTGTTGACCCAGATGAGG 240  
DB 419 AACAGATCTATGAGCGATTTGGCACTTCTGTTGATCTGTTGACCCAGATGAGG 478

QY 241 TCAGAAGCTTGAAGAAAGCATTCATTTATGACAAAACCTGACAGCATGGGCTG 300  
DB 479 TCAGAAGCTTGAAGAAAGCATTCATTTATGACAAAACCTGACAGCATGGGCTG 538

QY 301 GAGAAAGTTCACCTGAGCGAGTGAATACCCAGTGGGAGAGGAGAAAGATCAGCT 360  
DB 539 GAGAAAGTTCACCTGAGCGAGTGAATACCCAGTGGGAGAGGAGAAAGATCAGCT 598

QY 361 GTGATGCTGAGCGAGAAATCATAGATAGCATCTCTGCTTTGGCAGCAGATTGGG 420  
DB 599 GTGATGCTGAGCGAGAAATCATAGATAGCATCTCTGCTTTGGCAGCAGATTGGG 658

QY 421 ACTCTCTCAGAAAGCATTCAGCGAGAGTTGTTGTTGATCTCTTGGATGAAGCTGAG 480  
DB 659 ACTCTCTCAGAAAGCATTCAGCGAGAGTTGTTGTTGATCTCTTGGATGAAGCTGAG 718

QY 481 AGAAGGCTCTCAGAAAGCATTCAGCGAGAGTTGTTGTTAACAACCTTATCAACTAC 540  
DB 719 AGAAGGCTCTCAGAAAGCATTCAGCGAGAGTTGTTGTTAACAACCTTATCAACTAC 778

QY 541 TCAAGACGCTGCAATACGAAACGAGGGGCGGTGAAGCTGCCAAGGTGGGGCTTTG 600

DB 779 TCAAGACGCTGCAATACGAAACGAGGGGCGGTGAAGCTGCCAAGTTGGGGCTTTG 838  
QY 601 GCATCTCTATTCGATCCGTCGTCCTCTCTCATCTACAGTCTCCACAGGATTCAG 660  
DB 839 GCATCTCTATTCGATCCGTCGTCCTCTCTCATCTACAGTCTCCACAGGATTCAG 898

QY 661 GAATACAGGATGGCGTCCCAAAATTCACAGCCCTGTATTACGTGGAAGATGACAA 720  
DB 899 GAATACAGGATGGCGTCCCAAAATTCACAGCCCTGTATTACGTGGAAGATGACAA 958

QY 721 ATGATGTCAGAAATGCTTCTCATGGGATCAAAAT 756  
DB 959 ATGATGTCAGAAATGCTTCTCATGGGATCAAAAT 994

RESULT 14  
BX355939/c 996 bp mRNA linear EST 05-MAY-2003  
LOCUS  
DEFINITION BX355939 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA  
clone CSOD1006Y115 3-PRIME, mRNA sequence.  
ACCESSION BX355939  
VERSION BX355939.1 GI:30382027  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Genoscope -- Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4663.f For more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CSOD1006AB08NP1cluster=4663.f. Contact :  
Peng Liang Email : filiang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CSOD1006AB08NP1.

FEATURES  
SOURCE  
1. 996  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSOD1006Y115"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT  
270 a 231 c 223 g 269 t 3 others

Query Match 51.3%; Score 726.6; DB 13; Length 996;  
Best Local Similarity 99.9%; Pred. No. 2e-201; Indels 0; Gaps 0;  
Matches 726; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 690 AACAGCTGTATTTCGCTGAGATGCAAGAAATGATTCAGAAATGCTTCTCATGGAT 749  
DB 996 AACAGCTGTATTTCGCTGAGATGCAAGAAATGATTCAGAAATGCTTCTCATGGAT 937

QY 750 CAAATATGCTATTCAGCTTAAAGATGGGGCAAGACCTTACCGAGTCTGATTCCTTCAA 809  
DB 936 CAAATATGCTATTCAGCTTAAAGATGGGGCAAGACCTTACCGAGTCTGATTCCTTCAA 877

QY 810 CACTGTAGCAGAGATCACTGGGACCAATATCCAGAAACAGTTGTACTGTCTAGTGACA 869  
DB 876 CACTGTAGCAGAGATCACTGGGACCAATATCCAGAAACAGTTGTACTGTCTAGTGACA 817

```

Oy 870 TCTGACAGCTGGAGTGTGGCAGGAGTCCATGATGATGGGAGCCTTATATTC
Db 816 TCTGACAGCTGGAGTGTGGCAGGAGTCCATGATGATGGGAGCCTTATATTC
Oy 930 ATGGAGACACTCTCACTTATTAAGATCTTGGGCTGGCTCCAAAGAGACTCTGCGCT
Db 756 ATGGAGACACTCTCACTTATTAAGATCTTGGGCTGGCTCCAAAGAGACTCTGCGCT
Oy 990 GGTGCTCTGAGTGCAGAAAGAACAGGTGAGATTTGGTCTTCCAGTATTAATCACTTACA
Db 696 GGTGCTCTGAGTGCAGAAAGAACAGGTGAGATTTGGTCTTCCAGTATTAATCACTTACA
Oy 1050 CAAGGTAATATATTTCCAACTACAGTCTGGTATGAGTCTGACGAGAACCTTCTTACC
Db 636 CAAGGTAATATATTTCCAACTACAGTCTGGTATGAGTCTGACGAGAACCTTCTTACC
Oy 1110 CACTGGGCTGCATTCATCTGAGCAGTGAAGGCGCAGGCGCATCATGAGAGATTATGAG
Db 576 CACTGGGCTGCATTCATCTGAGCAGTGAAGGCGCAGGCGCATCATGAGAGATTATGAG
Oy 1170 CTTGCTGCAGCCCTTCATATATCAAGTCTTGAGCCATGAGAAAGGACAGATCA
Db 516 CTTGCTGCAGCCCTTCATATATCAAGTCTTGAGCCATGAGAAAGGACAGATCA
Oy 1230 CTTTGGATCCAGGCTGAGGCTGAGCCAGTCACTTGAATGACTTATACAGTATTT
Db 466 CTTTGGATCCAGGCTGAGGCTGAGCCAGTCACTTGAATGACTTATACAGTATTT
Oy 1290 CTTTTCATCATCTCCACAGGAGACACATGATCTGTATGATCCAAAGAGATGATGT
Db 396 CTTTTCATCATCTCCACAGGAGACACATGATCTGTATGATCCAAAGAGATGATGT
Oy 1350 TGTCTCTGCTGTTTGGGCTGTTTCTTATGTTGTGACACATGAGAAAGATCTGCC
Db 336 TGTCTCTGCTGTTTGGGCTGTTTCTTATGTTGTGACACATGAGAAAGATCTGCC
Oy 1410 TAGGTCC 1416
Db 276 TAGGTCC 270

```

```

RESULT 15
CB958693 739 bp mRNA linear EST 29-APR-2003
LOCUS AGNC00937.13778385 NIH_MGC_184 Homo sapiens cDNA clone
DEFINITION IMAGE:30351744 5', mRNA sequence.
ACCESSION CB958693
VERSION CB958693.1 GI:30214809
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 739)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palokovits
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILN at:
http://image.llnl.gov
Plate: NDCM149 row: m column: 01
High quality sequence stop: 572.
Location/Qualifiers
1..739
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

FEATURES
Source
1..739
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db xref="taxon:9606"
/clone="IMAGE:30351744"
/lab host="DH10B (TI phase-resistant)"
/clone_lib="NIH_MGC_184"
/note="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site_1:
Site1 (ggccattatggcc); Site_2: Site1 (ggccctccggcc);
Library is oligo-dT primed and directionally cloned. cDNA
was prepared from a glandular pool of tissues from thyroid,
parathyroid, adrenal, cortex and pineal gland. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTTAGC-3', and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGCGCATGATG-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.38
kb (range 0.60-3.5 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC Library."
BASE COUNT 192 a 171 c 195 g 181 t
ORIGIN
Query Match 50.8%; Score 719.2; DB 14; Length 739;
Best Local Similarity 99.6%; Pred. No. 2.6e-199;
Matches 721; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 588 GGTGGGGCTTTGGCATCTCATTCATCCGTCGCTCTTCTCATCTACAGTCTCA
Db 3 GGTGGGGCTTTGGCATCTCATTCATCCGTCGCTCTTCTCATCTACAGTCTCA
Oy 648 CACAGTATTCAGGAATACAGAGATGCGTCCCAAAATTCACAGCTGTATAGGT
Db 63 CACAGTATTCAGGAATACAGAGATGCGTCCCAAAATTCACAGCTGTATAGGT
Oy 708 GGAAGATGCAGAAATGATGTCAGAAATGGCTTCTCATGGGATCAAAATGTCAATCACT
Db 123 GGAAGATGCAGAAATGATGTCAGAAATGGCTTCTCATGGGATCAAAATGTCAATCACT
Oy 768 AAAGATGGGGCAAGACCTACCCAGATGATTCCTTCAACACTGATGAGAGATCAC
Db 183 AAAGATGGGGCAAGACCTACCCAGATGATTCCTTCAACACTGATGAGAGATCAC
Oy 828 TGGAGCAAAATATCAAGACAGTTGATCTGTCATGACATTTGACACAGCTGGATGT
Db 243 TGGAGCAAAATATCAAGACAGTTGATCTGTCATGACATTTGACACAGCTGGATGT
Oy 888 TGGGCAAGGTCATGATGATGAGGCGTGGAGCCTTATATCAAGGAGACTCTCACT
Db 303 TGGGCAAGGTCATGATGATGAGGCGTGGAGCCTTATATCAAGGAGACTCTCACT
Oy 948 TATTAAAGATCTTGGGCTGGCTCCAAAGAGACTCTGCGGTGCTGTGACATGACAGA
Db 363 TATTAAAGATCTTGGGCTGGCTCCAAAGAGACTCTGCGGTGCTGTGACATGACAGA
Oy 1008 AGAACAAAGTGAAGTGTGGCTTTCAGATTTATCACTACACAGGTAATATTTCAA
Db 423 AGAACAAAGTGAAGTGTGGCTTTCAGATTTATCACTACACAGGTAATATTTCAA
Oy 1068 CTAAGTCTGGTATGAGAGTCTGACGAGAACTTCTTCCACCTGGGCTGCAATTCAC
Db 483 CTAAGTCTGGTATGAGAGTCTGACGAGAACTTCTTCCACCTGGGCTGCAATTCAC
Oy 1128 TGGCAGTGAAGAGCCAGGCGCATCATGAGAGGTTATGAGCTGTGACGCCCTCAA
Db 543 TGGCAGTGAAGAGCCAGGCGCATCATGAGAGGTTATGAGCTGTGACGCCCTCAA
Oy 1188 TATCACTCAAGTCTGAGCCATGAGAGAGGACAGACATCACTTTGGATCCAGCTGG
Db 603 TATCACTCAAGTCTGAGCCATGAGAGAGGACAGACATCACTTTGGATCCAGCTGG
Oy 1248 AGTCCCTGAGCCAGCTACTGATGATGATCAATATTTCTTCCATCACTCCCA
Db 663 AGTCCCTGAGCCAGCTACTGATGATGATCAATATTTCTTCCATCACTCCCA
Oy 1308 CGGA 1311

```

Db 723 CGGA 726

Search completed: December 22, 2003, 23:11:28  
Job time : 2776.81 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:59:40 ; Search time 93.5249 Seconds  
(without alignments)  
6682.695 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_99\_1514  
Perfect score: 1416  
Sequence: 1 ATGAATTCCTATCTTCG.....AAGAATGCTGCTAGGTCC 1416

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTCUTS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1412.8	99.8	1863	US-09-482-273-28	Sequence 28, Appl
2	772.4	54.5	1134	US-09-482-273-95	Sequence 95, Appl
3	49.2	3.5	7218	US-08-232-463-14	Sequence 14, Appl
4	37	2.6	2049	US-08-268-797-1	Sequence 1, Appl
5	37	2.6	2049	PCT-US95-08414-1	Sequence 1, Appl
6	37	2.6	2691	US-07-878-960-1	Sequence 1, Appl
7	35.2	2.5	2373	US-08-980-080-1	Sequence 1, Appl
8	35.2	2.5	3003	US-08-434-730-15	Sequence 15, Appl
9	34.6	2.4	4086	US-09-220-132-56	Sequence 56, Appl
10	34.6	2.4	4334	US-09-620-312D-88	Sequence 88, Appl
11	34.4	2.4	738	US-08-738-462-1	Sequence 1, Appl
12	34.4	2.4	738	PCT-US94-07587-1	Sequence 1, Appl
13	34	2.4	2106	US-09-252-991A-7477	Sequence 7477, Ap
14	34	2.4	2112	US-09-252-991A-7334	Sequence 7334, Ap
15	34	2.4	2187	US-09-252-991A-7167	Sequence 7167, Ap
16	33.4	2.4	2201	US-09-330-970-2	Sequence 2, Appl
17	33.4	2.4	2201	US-09-330-970-4	Sequence 4, Appl
18	33.2	2.3	2443	US-08-745-934-2	Sequence 2, Appl
19	33.2	2.3	9763	US-08-973-273-1	Sequence 1, Appl
20	33	2.3	1165	US-09-023-942A-28	Sequence 28, Appl
21	33	2.3	1443	US-09-328-352-2934	Sequence 2934, Ap
22	33	2.3	4791	US-09-328-352-1416	Sequence 1416, Ap
23	32.4	2.3	3483	US-09-620-312D-658	Sequence 658, App
24	32.2	2.3	272	US-09-313-294A-3342	Sequence 3342, App
25	32.2	2.3	1650	US-09-220-132-103	Sequence 103, App
26	32.2	2.3	2220	US-08-97-251-1	Sequence 1, Appl
27	32	2.3	4631	US-09-620-312D-164	Sequence 164, App

28	31.8	2.2	615	US-09-107-532A-1080	Sequence 1080, Ap
29	31.6	2.2	949	US-09-221-017B-594	Sequence 594, App
c	30	31.6	28473	US-08-961-527-83	Sequence 83, Appl
31	31.4	2.2	2242	US-09-400-742-1	Sequence 1, Appl
32	31.4	2.2	2242	US-08-618-651A-1	Sequence 1, Appl
33	31.4	2.2	2242	US-09-215-252-1	Sequence 1, Appl
34	31.2	2.2	1712	US-08-809-326A-8	Sequence 8, Appl
35	31.2	2.2	1712	US-09-689-913A-8	Sequence 8, Appl
36	31.2	2.2	1712	US-09-689-913A-8	Sequence 8, Appl
37	31.2	2.2	1712	US-09-689-913A-8	Sequence 8, Appl
38	31.2	2.2	1927	US-09-689-914A-13	Sequence 13, Appl
39	31.2	2.2	1927	US-09-689-913A-13	Sequence 13, Appl
40	31.2	2.2	1927	US-09-689-913A-13	Sequence 13, Appl
41	31.2	2.2	1954	US-08-809-326A-13	Sequence 13, Appl
c	42	31	280	US-08-783-853A-34	Sequence 34, Appl
c	43	31	280	US-09-344-050-34	Sequence 34, Appl
c	44	31	321	US-08-783-853A-43	Sequence 43, Appl
c	45	31	321	US-09-344-050-43	Sequence 43, Appl

ALIGNMENTS

RESULT 1  
US-09-482-273-28  
Sequence 28, Application US/09482273.  
Patent No. 6534631  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 71 Human Secreted Proteins  
FILE REFERENCE: P2030P1  
CURRENT APPLICATION NUMBER: US/09/482,273  
CURRENT FILING DATE: 2000-01-13  
EARLIER APPLICATION NUMBER: PCT/US99/15849  
EARLIER FILING DATE: 1999-07-14  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/092,922  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/092,956  
EARLIER FILING DATE: 1998-07-15  
NUMBER OF SEQ ID NOS: 267  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 28  
LENGTH: 1863  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-482-273-28  
Query Match 99.8%; Score 1412.8; DB 4; Length 1863;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
c  
1 ATGAATTCCTATCTTCGATTTTGGTGGTTCACCTTTATCCCTGCTGGG 60  
ATGAATTCCTATCTTCGATTTTGGTGGTTCACCTTTATCCCTGCTGGG 158  
61 AAAGCTATGCAAGATGCGATCTCTAGAGACCTTTGAAGAAATTAAGAATA 120  
AAAGCTATGCAAGATGCGATCTCTAGAGACCTTTGAAGAAATTAAGAATA 218  
159 AAAGCTATGCAAGATGCGATCTCTAGAGACCTTTGAAGAAATTAAGAATA 218  
121 GCCAGCTGGAGATGTTGCTTAAGCAATCACTCAAGCTGTTTATGTAAGCCAG 180  
GCCAGCTGGAGATGTTGCTTAAGCAATCACTCAAGCTGTTTATGTAAGCCAG 278  
219 GCCAGCTGGAGATGTTGCTTAAGCAATCACTCAAGCTGTTTATGTAAGCCAG 278  
c  
181 AACAGATCTATGAGCGATGGCACTTCTGTTGATATCTGTGACCGAGATGGTGC 240  
AACAGATCTATGAGCGATGGCACTTCTGTTGATATCTGTGACCGAGATGGTGC 338  
279 AACAGATCTATGAGCGATGGCACTTCTGTTGATATCTGTGACCGAGATGGTGC 338  
c  
241 TCCAAGAACTTGAAGAAAGCCATCAATTAATGTAACAAACCTGACGAGATGGCTG 300  
TCCAAGAACTTGAAGAAAGCCATCAATTAATGTAACAAACCTGACGAGATGGCTG 398  
Db

301 GAGAAAGTTACCTGAGCCAGTGAATACCCACTGAGAGAGGAGAAATCAAGCT 360  
399 GAGAAAGTTACCTGAGCCAGTGAATACCCACTGAGAGAGGAGAAATCAAGCT 458  
361 GTGATGCTGAGCCAGAAATTCATAGATACCATCTGGGCTTTGGCAGACATTTGG 420  
459 GTGATGCTGAGCCAGAAATTCATAGATACCATCTGGGCTTTGGCAGACATTTGG 518  
421 ACTCTCCAGAAAGGATTAACAGCAAGATTTCTGGTGGAGACCTCTTTGAGAACTGAG 480  
519 ACTCTCCAGAAAGGATTAACAGCAAGATTTCTGGTGGAGACCTCTTTGAGAACTGAG 578  
481 AGAAGGCTCAGAAAGCAAGAGGAGAGATTTGTTTAAACCACTTACATCACTAC 540  
579 AGAAGGCTCAGAAAGCAAGAGGAGAGATTTGTTTAAACCACTTACATCACTAC 638  
541 TCAAGACGCTGCAATACCGAACGCGGGCGGTGGAAGCTGCCAGGTGGGGCTTTG 600  
639 TCAAGACGCTGCAATACCGAACGCGGGCGGTGGAAGCTGCCAGGTGGGGCTTTG 698  
601 GCATCTCTCATTCGATCCGCTGGCCCTCTCCATCTACAGTCCCTCAGACAGTATTCAG 660  
699 GCATCTCTCATTCGATCCGCTGGCCCTCTCTCCATCTACAGTCCCTCAGACAGTATTCAG 758  
661 GAATACCAAGATGGCGTCCCAAAATTCACACAGCTGTATTACGCTGGAAGATGCA 720  
759 GAATACCAAGATGGCGTCCCAAAATTCACACAGCTGTATTACGCTGGAAGATGCA 818  
721 ATGATGTCAGAAATGGCTTCTCATAGGATCAAAATTTGCTATTACGCTTAAAGATGG 780  
819 ATGATGTCAGAAATGGCTTCTCATAGGATCAAAATTTGCTATTACGCTTAAAGATGG 878  
781 AAGACCTACCAAGATGCTGATCTCTTCAACCTGTAGAGAGATTCACGCGAGCAAAAT 840  
875 AAGACCTACCAAGATGCTGATCTCTTCAACCTGTAGAGAGATTCACGCGAGCAAAAT 938  
841 CCAGAAACAGTTGTAAGTGTAGTGAATGATGCACTGAGCACTGGAGATTTGGCGAGTGGC 900  
935 CCAGAAACAGTTGTAAGTGTAGTGAATGATGCACTGAGCACTGGAGATTTGGCGAGTGGC 998  
901 ATGATGTCAGAAATGGCTTCTCATAGGATCAAAATTTGCTATTACGCTTAAAGATGG 960  
999 ATGATGTCAGAAATGGCTTCTCATAGGATCAAAATTTGCTATTACGCTTAAAGATGG 1058  
961 GGGCGCTGCAAGAGAGACTCTGGGCTGGTGGCTGAGCACTGAGCAAGAAACAGTGGGA 1020  
1059 GGGCGCTGCAAGAGAGACTCTGGGCTGGTGGCTGAGCACTGAGCAAGAAACAGTGGGA 1118  
1021 GTTGGTGGCTTCCAGATATTATCAAGTGAATTAATTTCCAACTACAGTCTGGTG 1080  
1119 GTTGGTGGCTTCCAGATATTATCAAGTGAATTAATTTCCAACTACAGTCTGGTG 1178  
1081 ATGAGATCTGACGAGAACTTCTTACCACTGGGCTGCAATTCATCTGGAGTGA 1140  
1179 ATGAGATCTGACGAGAACTTCTTACCACTGGGCTGCAATTCATCTGGAGTGA 1238  
1141 GCCAGGCGCAATCAGAGAGAGTATAGCTGTGAGCCCTCCAAATATCACTCAGGTC 1200  
1239 GCCAGGCGCAATCAGAGAGAGTATAGCTGTGAGCCCTCCAAATATCACTCAGGTC 1298  
1201 CTGAGCCATGAGAGAGAGAGAGATCACTTTTGGATCCAAAGCTGAGTCCCTGAGCC 1260  
1299 CTGAGCCATGAGAGAGAGAGAGATCACTTTTGGATCCAAAGCTGAGTCCCTGAGCC 1358  
1261 AGTCTACTGATGATCTTATACAGTATTTCTTCTTCATCACTCCACGAGAGACATG 1320  
1359 AGTCTACTGATGATCTTATACAGTATTTCTTCTTCATCACTCCACGAGAGACATG 1418  
1321 ACTGTCATGATCCAAAGAGATGATGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1380  
1419 ACTGTCATGATCCAAAGAGATGATGCTGCTGCTGTTGGGCTGTTGTTCTTAT 1478  
1381 GTTGTTCAGACATGAGAGAAATGCTGCTAGTCC 1416

Db 1479 GTTGTTCAGACATGAGAGAAATGCTGCTAGTCC 1514  
RESULT 2  
US-09-482-273-95  
Sequence 95, Application US/09482273  
Patent No. 6534631  
GENERAL INFORMATION:  
APPLICANT: Roese et al.  
TITLE OF INVENTION: 71 Human Secreted Proteins  
FILE REFERENCE: P2030P1  
CURRENT APPLICATION NUMBER: US/09/482,273  
EARLIER FILING DATE: 2000-01-13  
EARLIER APPLICATION NUMBER: PCT/US99/15849  
EARLIER FILING DATE: 1999-07-14  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/092,922  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/092,956  
NUMBER OF SEQ ID NOS: 267  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 95  
LENGTH: 1134  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-482-273-95  
Query Match 54.5%; Score 772.4; DB 4; Length 1134;  
Best Local Similarity 99.7%; Pred. No. 1e-233;  
Matches 784; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 631 TCATCTACAGTCTCTCAACAGTATTCAGAGATCCAGAGATGCGTCCCAATTCGA 690  
Db 1 TCATCTACAGTCTCTCAACAGTATTCAGAGATCCAGAGATGCGTCCCAATTCGA 60  
QY 691 ACAGCCGTATTCAGGAGAGATTCAGAAATGATGCAAGATGCTTCTCATAGGATTC 750  
Db 61 ACAGCCGTATTCAGGAGAGATTCAGAAATGATGCAAGATGCTTCTCATAGGATTC 120  
QY 751 AAAATGTCATTCAGCTTAAAGATGGGGCAAGACCTACCATCTGATTCCTTCAC 810  
Db 121 AAAATGTCATTCAGCTTAAAGATGGGGCAAGACCTACCATCTGATTCCTTCAC 180  
QY 811 ACTGACAGAGATCACTGGAGCAATATTCAGAACAGTGTACTGTGACATGAT 870  
Db 181 ACTGACAGAGATCACTGGAGCAATATTCAGAACAGTGTACTGTGACATGAT 240  
QY 871 CTGACAGCTGGAGATGTTGGGCAAGGCTGCATGATGATGAGCGGTGAGCTTATATCA 930  
Db 241 CTGACAGCTGGAGATGTTGGGCAAGGCTGCATGATGATGAGCGGTGAGCTTATATCA 300  
QY 931 TGGAGACATCTCATCTTATTAAGATCTTGGCTGCTGCTCAAGAGAGATCTGCGCTG 990  
Db 301 TGGAGACATCTCATCTTATTAAGATCTTGGCTGCTGCTCAAGAGAGATCTGCGCTG 360  
QY 991 GTGCTGGAAGTGGAGAGAGAGAGAGTGTGCTTCCAGTATTCAGTTAC 1050  
Db 361 GTGCTGGAAGTGGAGAGAGAGAGAGTGTGCTTCCAGTATTCAGTTAC 420  
QY 1051 AAGGTAAATTTTCCACTACATCTGATGATGATGATGATGATGATGATGATGATGAT 1110  
Db 421 AAGGTAAATTTTCCACTACATCTGATGATGATGATGATGATGATGATGATGATGAT 480  
QY 1111 ACTGGGCTGCAATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1170  
Db 481 ACTGGGCTGCAATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 539  
QY 1171 CTGCTGAGCCCTCAATATCACTCAGGCTCTGAGCCATGAGAGAGAGAGAGATCAAC 1230  
Db 540 CTGCTGAGCCCTCAATATCACTCAGGCTCTGAGCCATGAGAGAGAGAGAGATCAAC 599



49 CTGTGCTCTGGGAAGCTATATGC

US-08-268-797-1

49 CTGTCCTCTGGGAAAGCTATATGCAAGATGGCATCTCTAAGAGGACTTTTGAGCAATA 108



Dh 1947 GCTGACAGACATTGGTATGATGACATGACACCGCATTTGGCCATGATGTCAGCTCGG 1888  
Qy 1257 AGCCAGTCTACTGATGATGATTAATCAAGTATTTCTTCATGACTCCGACGAGACAC 1316  
Dh 1887 CAACAGGCTCTTTTGAACACTACACCATGTTTTCAGACTACCTTCAGCTTGAC 1828  
Qy 1317 CATGACTGTCAATGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGCTGTGTTTC 1376  
Dh 1827 CTGGAGAGACTTAAAGCCGACACAGGCGCCGATGCTCCGCTAACAGATTTCAATCAC 1768  
Qy 1377 TTAATGTTGT 1385  
Dh 1767 CAATGTGCT 1759

## RESULT 7

US-08-980-080-1  
Sequence 1, Application US/08980080  
Patent No. 6312941  
GENERAL INFORMATION:  
APPLICANT: CARTER-SU, CHRISTIN  
APPLICANT: RUI LIANG-YOU  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING  
TITLE OF INVENTION: SIGNALING PATHWAY AGONISTS AND ANTAGONISTS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MEDLEN & CARROLL, LLP  
STREET: 220 MONTGOMERY STREET, SUITE 2200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/980,080  
FILING DATE: 26-NOV-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CARROLL, PETER G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: UM-03036  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2373 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-980-080-1

Query Match 2.5%; Score 35.2; DB 4; Length 2373;

Best Local Similarity 44.4%; Pred. No. 1.3;

Matches 142; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

Dh 843 AGAAGAGTTGTATGCTGATGATGACATCTGGACAGCTGGGATGTTGGCAGGGTCCAT 902  
Qy 729 AGCAGAGATGATACAGAGAGAGAGCTGCTGATGTTTCATGAGGGCTGAAGAGGCTGCC 788  
Dh 903 GGATGATGCGCGTGGAGCTTTATATCATGGGAGCACTCTCACTTATTAAGATCTTGG 962  
Qy 789 TGACCCAGCAGAGATAGTGTGAGAGAGGGGAGCTGGCTGACCTCGGAGAGAGAGG 848  
Dh 963 GCTGCTCAAAAGAGACTCTGCGGCTGCTGCTGCTGCACTGCAAGAAACAAGGTGAGT 1022  
Qy 849 GCAGCTCAAGTGGCAGAAATGTGATTAATCTGCTCCGAGAGTGAAGAGAGAGAGG 908

Qy 1023 TGTGCTTCAGTATATATGATTAACAAGGTAATATTTCCAACTACAGTCTGTGAT 1082  
Dh 909 AAGTGGTTGAGATGTTTGTATACACCCAGAGCATCCCGCGTGTAGCATTCCTG 968  
Qy 1083 GAGCTGAGAGCAGGAACCTTTTACCACTGGGCTGCATTTCACTGCACTGAAAAGC 1142  
Dh 969 TTCTACTATTACTGATGATGCTGACAGCAGCAGCCCTGGAGATGCTGACAGGAGAAC 1028  
Qy 1143 CAGGCGCATCATGAGAGAGG 1162  
Dh 1029 GTTGTGTGTTAAGTAGAAG 1048

## RESULT 8

US-08-434-730-15  
Sequence 15, Application US/08434730  
Patent No. 5637463  
GENERAL INFORMATION:  
APPLICANT: Dalton, Stephen  
APPLICANT: Kochan, Jarema P  
APPLICANT: Osborne, Mark A  
TITLE OF INVENTION: METHOD TO DETECT PROTEIN-PROTEIN  
TITLE OF INVENTION: INTERACTIONS  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hoffmann-La Roche Inc  
STREET: 340 Kingsland Street  
CITY: Nutley  
STATE: NJ  
COUNTRY: USA  
ZIP: 07110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/434,730  
FILING DATE: 04-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Semionow, Raina  
REGISTRATION NUMBER: 39022  
REFERENCE/DOCKET NUMBER: 9069  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201)235-4391  
TELEFAX: (201)235-2363  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3003 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
US-08-434-730-15

Query Match 2.5%; Score 35.2; DB 1; Length 3003;

Best Local Similarity 44.4%; Pred. No. 1.5;

Matches 142; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

Dh 843 AGAAGAGTTGTATGCTGATGATGACATCTGGACAGCTGGGATGTTGGCAGGGTCCAT 902  
Qy 1072 AGCAGAGATGATACAGAGAGAGAGCTGCTGATGTTTCATGAGGGCTGAAGAGGCTGCC 1131  
Dh 903 GGATGATGCGCGTGGAGCTTTATATCATGGGAGCACTCTCACTTATTAAGATCTTGG 962  
Qy 1132 TGACCCAGCAGAGATAGTGTGAGAGAGGGGAGCTGGCTGACCTCGGAGAGAGAGG 1191  
Dh 963 GCTGCTCAAAAGAGACTCTGCGGCTGCTGCTGCTGCACTGCAAGAAACAAGGTGAGT 1022  
Qy 1192 GCAGCTCAAGTGGCAGAAATGTGATTAATCTGCTCCGAGAGTGAAGAGAGAGAGG 1251



STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 61..738  
US-08-738-462-1

Query Match 2.4%; Score 34.4; DB 2; Length 738;  
Best Local Similarity 48.4%; Pred. No. 1.1;  
Matches 124; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

QY 696 CTGTATTACGGTGAAGATGCAGAAATGATGTCAAGATGGCTTCTCATGGGATCAAAAT 755  
DB 261 CTGTATTACGGTGAAGATGCAGAAATGATGTCAAGATGGCTTCTCATGGGATCAAAAT 202  
QY 756 TGTCAATTCAGCTAAGATGGGGCAAGACCTTACCAGAT-ACGTATTCCTTCAACACTG 814  
DB 201 CCCAGCACACTTAAGTTTCAGGGCAAAATGATGCCAAGTGAMCCCTGACAGACATCCG 142  
QY 815 TAGCAGATCATCTGGAGCAAAATATCCAGAAAGTTGTACTGTGATGACATCTGG 874  
DB 141 AGAAGAGATGAAGGTAAAGAACTTCTGGGCTGAGTCTGTGGTAAAGTTTAACTAGC 82  
QY 875 ACAGCTGGATGTTGGCAGGGTGCCATGATGATGCGGTGAGCCCTTATATCATGGG 934  
DB 81 ATATATTGCTTGTGTCATGCGACAGAGAGAGGTGGATTAACCTGTGTGAAGA 22  
QY 935 AAGCACTCTCACTTAT 950  
DB 21 ACAACTCTCAATTTCT 6

## RESULT 12

PCT-US94-07587-1/c  
Sequence 1, Application PC/TUS9407587  
GENERAL INFORMATION:  
APPLICANT: Schering Corp.  
TITLE OF INVENTION: PURIFIED MAMMALIAN NK ANTIGENS AND  
TITLE OF INVENTION: RELATED REAGENTS  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schering Corp.  
STREET: One Giraldia Farms  
City: Madison  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07940  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.1  
SOFTWARE: Microsoft Word 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/07587  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lunn, Paul G.  
REGISTRATION NUMBER: 32,743  
REFERENCE/DOCKET NUMBER: DX0397K  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-822-7255  
TELEFAX: 201-822-7039  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 738 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS

LOCATION: 61..738  
PCT-US94-07587-1

Query Match 2.4%; Score 34.4; DB 5; Length 738;  
Best Local Similarity 48.4%; Pred. No. 1.1;  
Matches 124; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

QY 696 CTGTATTACGGTGAAGATGCAGAAATGATGTCAAGATGGCTTCTCATGGGATCAAAAT 755  
DB 261 CTGTATTACGGTGAAGATGCAGAAATGATGTCAAGATGGCTTCTCATGGGATCAAAAT 202  
QY 756 TGTCAATTCAGCTAAGATGGGGCAAGACCTTACCAGAT-ACGTATTCCTTCAACACTG 814  
DB 201 CCCAGCACACTTAAGTTTCAGGGCAAAATGATGCCAAGTGAAACCTGACAGACATCCG 142  
QY 815 TAGCAGATCATCTGGAGCAAAATATCCAGAAAGTTGTACTGTGATGACATCTGG 874  
DB 141 AGAAGAGATGAAGGTAAAGAACTTCTGGGCTGAGTCTGTGGTAAAGTTTAACTAGC 82  
QY 875 ACAGCTGGATGTTGGCAGGGTGCCATGATGATGCGGTGAGCCCTTATATCATGGG 934  
DB 81 ATATATTGCTTGTGTCATGCGACAGAGAGAGGTGGATTAACCTGTGTGAAGA 22  
QY 935 AAGCACTCTCACTTAT 950  
DB 21 ACAACTCTCAATTTCT 6

## RESULT 13

US-09-252-991A-7477  
Sequence 7477, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 7477  
LENGTH: 2106  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7477

Query Match 2.4%; Score 34; DB 4; Length 2106;  
Best Local Similarity 54.9%; Pred. No. 2.8;  
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1042 CAGTTACACAAGGTAATATTTCCAACTACAGTCTGGTGAAGTGTGACGAGAAAC 1101  
DB 1181 CAGGAAGTGAAGACCTGATGCTTCTTCTGCTGTGTGAACCCGACGACGACAC 1240  
QY 1102 TTCTTACCACTGGGCTCAATTCATCTGCGAGTGAAGGCAAGGCCATCATGAGAG 1161  
DB 1241 GCCTTCTCAGGGTGAATCAAGTGGCGCGGGGAAATCGTTTCGGGACCTGGAAG 1300  
QY 1162 GT 1163  
DB 1301 CT 1302

## RESULT 14

US-09-252-991A-7334  
Sequence 7334, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 14:49:05 ; Search time 443.013 Seconds  
(without alignments)  
10650.710 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_99\_1514  
Perfect score: 1416  
Sequence: 1 ATGAATTCCTTATCTTCG.....AAGAAATGCTGCTAGTCC 1416

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2211978 seqs, 1666101734 residues

Total number of hits satisfying chosen parameters: 4423956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1416	100.0	1851	9	US-09-745-763-35
2	1412.8	99.8	1863	11	US-09-984-271-28
3	1107.2	78.2	1778	10	US-09-917-800A-505
4	772.4	54.5	1134	11	US-09-984-371-95
5	261.8	18.5	357	10	US-09-833-381-1929
6	246.2	17.4	427	10	US-09-833-381-1930
7	91	6.4	134	10	US-09-783-590-12141
8	60.2	4.3	65	13	US-09-908-975-25931
9	40.2	4.2	60	13	US-09-908-975-8960
10	42	3.0	2200	15	US-10-059-585-5
11	41	2.9	65	13	US-09-908-975-1202
12	39.4	2.8	484	11	US-09-918-995-30889
13	39.4	2.8	1363	15	US-10-232-484-5
14	39.4	2.8	2319	14	US-10-044-090-579
15	39.4	2.8	3322	13	US-09-814-353-20277

C	16	38.6	2.7	3997	15	US-10-198-846-12639	Sequence 12639, A
	17	38	2.7	457	13 <td>US-10-027-633-310889</td> <td>Sequence 310889,</td>	US-10-027-633-310889	Sequence 310889,
	18	38	2.7	457	14 <td>US-10-027-633-310889</td> <td>Sequence 310889,</td>	US-10-027-633-310889	Sequence 310889,
	19	37.8	2.7	999	15 <td>US-10-184-644-434</td> <td>Sequence 434, App</td>	US-10-184-644-434	Sequence 434, App
	20	37.8	2.7	999	15 <td>US-10-184-644-434</td> <td>Sequence 434, App</td>	US-10-184-644-434	Sequence 434, App
	21	37.8	2.6	502	15 <td>US-10-066-543-2028</td> <td>Sequence 2028, Ap</td>	US-10-066-543-2028	Sequence 2028, Ap
	22	37	2.6	548	15 <td>US-10-066-543-2222</td> <td>Sequence 2222, Ap</td>	US-10-066-543-2222	Sequence 2222, Ap
	23	37	2.6	568	10 <td>US-09-998-598-2545</td> <td>Sequence 2545, Ap</td>	US-09-998-598-2545	Sequence 2545, Ap
	24	37	2.6	2691	9 <td>US-09-922-217-121</td> <td>Sequence 121, App</td>	US-09-922-217-121	Sequence 121, App
	25	37	2.6	2691	10 <td>US-09-833-263-121</td> <td>Sequence 121, App</td>	US-09-833-263-121	Sequence 121, App
	26	37	2.6	2691	10 <td>US-09-880-107-2399</td> <td>Sequence 2399, Ap</td>	US-09-880-107-2399	Sequence 2399, Ap
	27	37	2.6	2691	13 <td>US-10-301-822-200</td> <td>Sequence 200, App</td>	US-10-301-822-200	Sequence 200, App
	28	37	2.6	2691	13 <td>US-09-873-367C-1009</td> <td>Sequence 1009, Ap</td>	US-09-873-367C-1009	Sequence 1009, Ap
	29	37	2.6	2691	14 <td>US-10-025-380-121</td> <td>Sequence 121, App</td>	US-10-025-380-121	Sequence 121, App
	30	37	2.6	2691	15 <td>US-10-102-524-1708</td> <td>Sequence 1708, Ap</td>	US-10-102-524-1708	Sequence 1708, Ap
	31	37	2.6	2782	14 <td>US-10-044-090-647</td> <td>Sequence 647, App</td>	US-10-044-090-647	Sequence 647, App
	32	37	2.6	2782	15 <td>US-10-208-408-16</td> <td>Sequence 16, App1</td>	US-10-208-408-16	Sequence 16, App1
	33	37	2.6	2782	15 <td>US-10-084-817-28</td> <td>Sequence 28, App1</td>	US-10-084-817-28	Sequence 28, App1
	34	37	2.6	3313	13 <td>US-09-814-353-19947</td> <td>Sequence 19947, A</td>	US-09-814-353-19947	Sequence 19947, A
	35	37	2.6	3320	15 <td>US-10-198-846-11549</td> <td>Sequence 11549, A</td>	US-10-198-846-11549	Sequence 11549, A
	36	36.4	2.6	5085	15 <td>US-10-198-846-9854</td> <td>Sequence 9854, Ap</td>	US-10-198-846-9854	Sequence 9854, Ap
	37	35.4	2.5	2574	13 <td>US-10-311-628-13</td> <td>Sequence 13, App1</td>	US-10-311-628-13	Sequence 13, App1
	38	35.2	2.5	1115	15 <td>US-10-184-644-440</td> <td>Sequence 440, App</td>	US-10-184-644-440	Sequence 440, App
	39	35.2	2.5	1115	15 <td>US-10-184-644-440</td> <td>Sequence 440, App</td>	US-10-184-644-440	Sequence 440, App
	40	35.2	2.5	715517	13 <td>US-10-027-633-53712</td> <td>Sequence 53712, A</td>	US-10-027-633-53712	Sequence 53712, A
	41	35.2	2.5	715517	14 <td>US-10-027-633-53712</td> <td>Sequence 53712, A</td>	US-10-027-633-53712	Sequence 53712, A
	42	35.2	2.4	416	11 <td>US-09-918-995-16569</td> <td>Sequence 16569, A</td>	US-09-918-995-16569	Sequence 16569, A
	43	34.6	2.4	1237	13 <td>US-10-027-632-199318</td> <td>Sequence 199318,</td>	US-10-027-632-199318	Sequence 199318,
	44	34.6	2.4	1237	14 <td>US-10-027-632-199318</td> <td>Sequence 199318,</td>	US-10-027-632-199318	Sequence 199318,
	45	34.6	2.4	1482	15 <td>US-10-060-036-66</td> <td>Sequence 66, App1</td>	US-10-060-036-66	Sequence 66, App1

#### ALIGNMENTS

RESULT 1  
US-09-745-763-35  
Sequence 35, Application US/09745763  
Patent No. US20020065394A1

GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
McCoy, John M.  
Lavallee, Edward R.  
Collins-Racie, Lisa A.  
Evans, Cheryl  
Merberg, David  
Treacy, Maurice  
Spaulding, Vikki

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM

NUMBER OF SEQUENCES: 219  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/745,763  
FILING DATE: 18-Jun-2000  
CLASSIFICATION: <unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Sprunger, Suzanne A.  
REGISTRATION NUMBER: 41,323  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8284  
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1851 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 35:  
US-09-745-763-35

Query Match 100.0%; Score 1416; DB 9; Length 1851;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAATTCCTTATCTTGCGATTTTTCGGTGGTGTACCTTTATACCTGTCGCTGGG 60
DB 99 ATGAATTCCTTATCTTGCGATTTTTCGGTGGTGTACCTTTATACCTGTCGCTGGG 158
QY 61 AAAGCTATATGCAAGATGCGATCTCTAAGAGACTTTTGAAGAAATAAAGAAATA 120
DB 159 AAAGCTATATGCAAGATGCGATCTCTAAGAGACTTTTGAAGAAATAAAGAAATA 218
QY 121 GCCAGCTGTGAGATGTTGCTTAAGCAATCATCAACTAGCTGTTTATGTAAGCCAG 180
DB 219 GCCAGCTGTGAGATGTTGCTTAAGCAATCATCAACTAGCTGTTTATGTAAGCCAG 278
QY 181 AACAGATCCTATGAGCGATGCGACTTCTGTTGTAAGCTGTTGAGACCGAGACTGAGTGC 240
DB 279 AACAGATCCTATGAGCGATGCGACTTCTGTTGTAAGCTGTTGAGACCGAGACTGAGTGC 338
QY 241 TCCAGAAACCTTAGAAAAAGCCATCCAAATATATGTAACAAACCTGAGAGATGGGCTG 300
DB 339 TCCAGAAACCTTAGAAAAAGCCATCCAAATATATGTAACAAACCTGAGAGATGGGCTG 398
QY 301 GAGAAAGTTTCACTTGGAGCGCAAGTATACCCCACTGGGAGAGGGAGAGATCACT 360
DB 399 GAGAAAGTTTCACTTGGAGCGCAAGTATACCCCACTGGGAGAGGGAGAGATCACT 458
QY 361 GTGATGCTGGAGCGCAAGATTCATAGATAGCCATCCTGGGCTTGGAGCGAGCTTGGG 420
DB 459 GTGATGCTGGAGCGCAAGATTCATAGATAGCCATCCTGGGCTTGGAGCGAGCTTGGG 518
QY 421 ACTCCTCCAGAGGAGATTACAGCAGAAAGTTCTGTGTGAGACTCTTTGATGAATGAG 480
DB 519 ACTCCTCCAGAGGAGATTACAGCAGAAAGTTCTGTGTGAGACTCTTTGATGAATGAG 578
QY 481 AGAAGGCGCTCAGAGAGAGGAGAGATGTTGTTTATTAACCAACTTACATCACTAC 540
DB 579 AGAAGGCGCTCAGAGAGAGGAGAGATGTTGTTTATTAACCAACTTACATCACTAC 638
QY 541 TCAAGGAGGCTGATACCGAAGCGAGGGGGGTGGAGAGCTGCCAAGGTGGGGCTTGG 600
DB 639 TCAAGGAGGCTGATACCGAAGCGAGGGGGGTGGAGAGCTGCCAAGGTGGGGCTTGG 698
QY 601 GCATCTCATTCGATCCGTCGCTCTTCTCCATCTACAGTCTCACAAGATATTCAG 660
DB 699 GCATCTCATTCGATCCGTCGCTCTTCTCCATCTACAGTCTCACAAGATATTCAG 758
QY 661 GAATACCAAGATGCGTGGCCCAAAATTCACAAGCCTGTATTACGTTGAGAGATCAGAA 720
DB 759 GAATACCAAGATGCGTGGCCCAAAATTCACAAGCCTGTATTACGTTGAGAGATCAGAA 818
QY 721 ATGATGTACAGAAATGGCTTCTCATAGGATCAAAATTCATTCACATTAAGATGGGGCA 780
DB 819 ATGATGTACAGAAATGGCTTCTCATAGGATCAAAATTCATTCACATTAAGATGGGGCA 878
QY 781 AAGACCTACCCAGATAGTATCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATAT 840
DB 879 AAGACCTACCCAGATAGTATCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATAT 938
QY 841 CCAGAACAGATTTGTAAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 900
DB 939 CCAGAACAGATTTGTAAGTGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 998
```

```
QY 901 ATGATGATGCGCGTGGAGACCTTTATATATGGAAGCACTCTCACTTATTAAGATCTT 960
DB 999 ATGATGATGCGCGTGGAGACCTTTATATATGGAAGCACTCTCACTTATTAAGATCTT 1058
QY 961 GGGCTGCGTCCAAAGAGAGACTCTGCGGCTGTGTCTGAGCTGAGAGAAACAAGTGA 1020
DB 1059 GGGCTGCGTCCAAAGAGAGACTCTGCGGCTGTGTCTGAGCTGAGAGAAACAAGTGA 1118
QY 1021 GTTGTGCTTCCAGATATATCACTTACCAAGATTAATTTTCCAACTACAGCTGTGTG 1080
DB 1119 GTTGTGCTTCCAGATATATCACTTACCAAGATTAATTTTCCAACTACAGCTGTGTG 1178
QY 1081 ATGAGTGTGACGAGAGAACCTTCTTACCACTGAGGCTGAATTCATGAGCAAGAAAG 1140
DB 1179 ATGAGTGTGACGAGAGAACCTTCTTACCACTGAGGCTGAATTCATGAGCAAGAAAG 1238
QY 1141 GCCAGGGCATCATGAGAGAGATTATGAGCTGTGAGAGCCCTCAATATCATCTCAGTGC 1200
DB 1239 GCCAGGGCATCATGAGAGAGATTATGAGCTGTGAGAGCCCTCAATATCATCTCAGTGC 1298
QY 1201 CTGAGCCATGAGAGAGAGAGAGAGATCACTTTTGAATCCAGCTGAGAGTCTGAGGCC 1260
DB 1299 CTGAGCCATGAGAGAGAGAGAGAGATCACTTTTGAATCCAGCTGAGAGTCTGAGGCC 1358
QY 1261 AGTCTACTGATGATGATTAACAAGTATTTCTTTCATCATCTGCCAGGAGACCATG 1320
DB 1359 AGTCTACTGATGATGATTAACAAGTATTTCTTTCATCATCTGCCAGGAGACCATG 1418
QY 1321 ACTGTATGATATCCAAAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1380
DB 1419 ACTGTATGATATCCAAAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1478
QY 1381 GTTGTGCAAGATGAGAGAAATGCTGCTTACGTTCC 1416
DB 1479 GTTGTGCAAGATGAGAGAAATGCTGCTTACGTTCC 1514
```

## RESULT 2

US-09-984-271-28  
Sequence 28, Application US/09984271  
Publication No. US20030040088A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 71 Human Secreted Proteins  
FILE REFERENCE: P2030P1  
CURRENT FILING DATE: US/09/984,271  
PRIOR APPLICATION NUMBER: 09/482,273  
PRIOR FILING DATE: 2000-01-13  
PRIOR APPLICATION NUMBER: PCT/US99/15849  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: 60/092,921  
PRIOR FILING DATE: 1998-07-15  
PRIOR APPLICATION NUMBER: 60/092,922  
PRIOR FILING DATE: 1998-07-15  
PRIOR APPLICATION NUMBER: 60/092,956  
NUMBER OF SEQ ID NOS: 267  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 28  
LENGTH: 1863  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-984-271-28

Query Match 99.8%; Score 1412.8; DB 11; Length 1863;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1414; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGAATTCCTTATCTTGCGATTTTTCGGTGGTGTACCTTTATACCTGTCGCTGGG 60
DB 99 ATGAATTCCTTATCTTGCGATTTTTCGGTGGTGTACCTTTATACCTGTCGCTGGG 158
```



QY	61	AAAGCTAATACGAAGAAATGGCATCTCTAAAGGACATTTTGAAGAAATTAAGAAATA	120
Db	159	AAAGCTAATACGAAGAAATGGCATCTCTAAAGGACATTTTGAAGAAATTAAGAAATA	218
QY	121	GCCAGCTGTGAGATGTGTCTAAAGCAATCAATCAACTAGCTGTTATAGTAAAGCCAG	180
Db	219	GCCAGCTGTGAGATGTGTCTAAAGCAATCAATCAACTAGCTGTTATAGTAAAGCCAG	278
QY	181	AACGATCTTATGACGATTGGCACTTGTGTTGATCTGTTGACCCAGACTGAGTGGC	240
Db	279	AACGATCTTATGACGATTGGCACTTGTGTTGATCTGTTGACCCAGACTGAGTGGC	338
QY	241	TCCAAAGAACCTTAAAGAAAGCCATCCAAATATATGTCACAAACCTGACACAGATGGCTG	300
Db	339	TCCAAAGAACCTTAAAGAAAGCCATCCAAATATATGTCACAAACCTGACACAGATGGCTG	398
QY	301	GAGAAAGTTCACTCGAGACAGTAGATATACCCCACTGGAGAGGGAGAAAGATCACT	360
Db	399	GAGAAAGTTCACTCGAGACAGTAGATATACCCCACTGGAGAGGGAGAAAGATCACT	458
QY	361	GTGATGCTGAGCCCAAGAAATTCATTAAGATGACATCTGGGTCTTGGACGACATTGGG	420
Db	459	GTGATGCTGAGCCCAAGAAATTCATTAAGATGACATCTGGGTCTTGGACGACATTGGG	518
QY	421	ACTCTCTCAGAAAGCATTCACGACAGAAATTCGTGGTGGTGAACCTCTTGGATGAAC	480
Db	519	ACTCTCTCAGAAAGCATTCACGACAGAAATTCGTGGTGGTGAACCTCTTGGATGAAC	578
QY	481	AGAAAGGCTCAGAAAGCAAGAGGGAGATGTGTTATTAACCAACTTACATCACTAC	540
Db	579	AGAAAGGCTCAGAAAGCAAGAGGGAGATGTGTTATTAACCAACTTACATCACTAC	638
QY	541	TCAAAGACGATGCATATACCGAACGACGAGGGCGGTGAAAGCTGCAGAGTGGGGCTTTG	600
Db	639	TCAAAGACGATGCATATACCGAACGACGAGGGCGGTGAAAGCTGCAGAGTGGGGCTTTG	698
QY	601	GCATCTCTCATTTCCATCCGTGGCTCTCTTCTTCATCTACAGTCTCTCACACAGTATTCAG	660
Db	699	GCATCTCTCATTTCCATCCGTGGCTCTCTTCTTCATCTACAGTCTCTCACACAGTATTCAG	758
QY	661	GAATACCAAGATGGCGTGGCCCAAAATTCACACAGCTGTATTAAGGTGAAAGATGCAGAA	720
Db	759	GAATACCAAGATGGCGTGGCCCAAAATTCACACAGCTGTATTAAGGTGAAAGATGCAGAA	818
QY	721	ATGATGTCAAAATATGGCTTCTCATGGGATCAAAATGTCAATTCAGCTAAAGATGGGGGCA	780
Db	819	ATGATGTCAAAATATGGCTTCTCATGGGATCAAAATGTCAATTCAGCTAAAGATGGGGGCA	878
QY	781	AAGACTACCCAGATACGATTCCTTCAACACTGACGACAGATCACTGGGAGCAAAATAT	840
Db	879	AAGACTACCCAGATACGATTCCTTCAACACTGACGACAGATCACTGGGAGCAAAATAT	938
QY	841	CCAGAAACAGTTGTACTGTGTCAGTGTGACATCTGACACACTGGGAATGTTGGGACGGGTGCC	900
Db	939	CCAGAAACAGTTGTACTGTGTCAGTGTGACATCTGACACACTGGGAATGTTGGGACGGGTGCC	998
QY	901	ATGATGTATGGCGGTGGAGCCCTTATATCATGGGAAAGACCTCTCACTTATTAAGATCTT	960
Db	999	ATGATGTATGGCGGTGGAGCCCTTATATCATGGGAAAGACCTCTCACTTATTAAGATCTT	1058
QY	961	GGGCTGCCTCCAAAGAGAGACTCTGGCGCTGGTGTCTGTGACTGCAAGAAACAAGTGTGA	1020
Db	1059	GGGCTGCCTCCAAAGAGAGACTCTGGCGCTGGTGTCTGTGACTGCAAGAAACAAGTGTGA	1118
QY	1021	GTGTGTGCTTCCAGATATTAATCAAGTTAAATATTTTCCAATCAAGTCTGTGTG	1080
Db	1119	GTGTGTGCTTCCAGATATTAATCAAGTTAAATATTTTCCAATCAAGTCTGTGTG	1178
QY	1081	ATGAGGTGTGACGCAAGAACTCTTACCACTGGGCTGCAAATTCACGTGGCAGTGAANAAG	1144
Db	1179	ATGAGGTGTGACGCAAGAACTCTTACCACTGGGCTGCAAATTCACGTGGCAGTGAANAAG	1238

[illegible]

```

Db      181  AAAGTATATACAGAGTGTGTTCTCAGCGAATTTCAAGAAATAAAGAAATA 240
Qy      121  GCCAGCTGAGAGTGTGCTTAACAGCATCACTAGCTGTTATAGTAAAGCCAG 180
Db      241  GCCAATATGAGAGTGTGCTTAACAGCATCACTAGCTGTTATAGTAAAGCCAG 300
Qy      181  AACAGATCTTAAGAGAGTGTGCACTTGTGTTATGTTGAGCCAGCTAGTGGC 240
Db      301  AACCGTGTATAGAGCTTTGGGACTTCTAGTTGTTACTGTTGGACCCAGATGATGGC 360
Qy      241  TCCAGAACTTAAGAAAGCCATCCAAATTAATGACCAAACTTGACAGCAAGTGGCTG 300
Db      361  TCTAAGAACTTAAGAAAGCTTAATCAATGATGACCAAACTTGACAGCAAGTGGCTG 420
Qy      301  GAGAAAGTTCACTGAGAGCCAGTGAATACCCCACTGGAGAGGGGAGGAATCACT 360
Db      421  GAAAACTGTCACCTGAGAGAGGTCAAGTAATCCTCACTGGGAGGGGAGGAATCTGCA 480
Qy      361  GTGATGCTGAGAGCCAAAGATTCAATAGATAGCCATCTGGGTCTTGGGAGAGCATTTGG 420
Db      481  GTGATGCTGAGAGCCAAAGATTCAATAGATAGCCATCTGGGTCTTGGGAGAGCATTTGG 540
Qy      421  ACTCTCTCAGAGAGCATTAACAGAGATTCGTGTGTGATGACCTCTTTGATGAATGTCAG 480
Db      541  ACTCTCTCAGAGAGCATTAACAGAGATTCGTGTGTGATGACCTCTTTGATGAATGTCAG 600
Qy      481  AGAAGGGCTTCAAGAGCAAGAGGAGAGTGTGTTTATTAACCAACTTATCACTAC 540
Db      601  AGAAGGGCATCAAGAGCAAGAGGAGAGTGTGTTTATTAACCAACTTATCACTAC 660
Qy      541  TCAAGAGCGTGCATTAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
Db      661  GGGAGAACTGTGTCAGTACCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
Qy      601  GCATCTCTCATTCATTCAGTCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
Db      721  GCATCTCTCATTCATTCAGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Qy      661  GATATCCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db      781  GATATCCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy      721  ATGATGTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db      841  ATGATGTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Qy      781  AAGACTTACCAAGATTAATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAATAT 840
Db      901  AAGACTTACCAAGATTAATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAATAT 960
Qy      841  CCAAGACAGGTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db      961  CCAAGACAGGTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Qy      901  ATGATGTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db      1021  CTGATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Qy      961  GGGCTGCTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
Db      1081  GGGCTGCTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Qy      1021  GTTGTGCTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Db      1141  GTTGTGCTTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Qy      1081  ATGAGTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db      1201  ATGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Qy      1141  GCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200

```

```

Db      1261  GCCAGGCTATCATGAGAGAGTCAAGTCTCTCGAACCCCTCAATATACCAAGTTC 1320
Qy      1201  CTGAGCATGAGAGAGGAGAGAGATCACTTTTGGATTCAGAGTGGAGTGGAGGCC 1260
Db      1321  TTTATGATGAGAGAGAGATGACATTAATCTTGTGATTCAGAGTGGAGTGGAGGCC 1380
Qy      1261  AGTCTACTGATGATGATTAATCAAGTATTTCTTCCATCATCTCCAGAGACCATG 1320
Db      1381  AGTCTGAGAGTATGATCTTGTACAGTATTTCTTTTCCATCATCTCCAGAGACCATG 1440
Qy      1321  ACTGTCATGATTCAGAGAGAGATGATGTTGCTGCTGCTGTTGGCTGTTGTTCTTAT 1380
Db      1441  ACTGTCATGATTCAGAGAGAGATGATGTTGCTGCTGCTGTTGGCTGTTGTTCTTAT 1500
Qy      1381  GTTGTGAGAGCATGAGAGAGAGTCTGCTAGGTC 1416
Db      1501  GTTGTGAGAGCATGAGAGAGAGTCTGCTAGGTC 1536

RESULT 4
US-09-984-271-95
; Sequence 95, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984,271
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482,273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092,956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 95
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-984-271-95

Query Match      54.5%; Score 772.4; DB 11; Length 1134;
Best Local Similarity 99.7%; Pred. No. 1.6e-240;
Matches 784; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```



```

; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16,2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12141
; LENGTH: 134
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (29)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (114)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (118)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (126)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-12141

Query Match      6.4%; Score 91; DB 10; Length 134;
Best Local Similarity 83.5%; Pred. No. 6.9e-19;
Matches 11; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

Oy      535 AACTACTAAGAGACGTCATATACGAAAGAGGCGGCGTGAAGTCCCAAGGTGGG 594
      |||
Db      2  ANCTACTAAGAGACGTCATATACGAAAGAGGCGGCGTGAAGTCCCAAGGTGGG 61

Oy      595 GCTTTGCAATCTCTCATTCG-ATCCGTGGCTCTTTCATCTAAGTCTCACAAG 653
      |||
Db      62  GCTTTGCAATCTTTATTTGGAATCGTGCCCTTTTTCATCTAAGTCTTAAANAG 121

Oy      654 TATTCAGGAATAC 666
      |||
Db      122 TTTNAGGATAC 134

RESULT 8
US-09-908-975-25931
; Sequence 25931, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSEMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLIC
; FILE REFERENCE: 36688-0005

```

```

; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25931
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-908-975-25931

Query Match      4.3%; Score 60.2; DB 13; Length 65;
Best Local Similarity 95.4%; Pred. No. 4.8e-09;
Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy      1333 CCAAGCAGATGATGTGCTGCTGTTGGAGCTGTCTTCTTATGTTGACAGAC 1392
      |||
Db      1  CCAAGCAGATGATGTGCTGCTGCTGTTATGGCTGTCTTATGTTGACAGAC 60

Oy      1393 ATGGA 1397
      |||
Db      61 ATGGA 65

RESULT 9
US-09-908-975-8960
; Sequence 8960, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSEMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLIC
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8960
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-908-975-8960

Query Match      4.2%; Score 60; DB 13; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.3e-09;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      152 TCAACTAGCTGTTTATGTTAAGCCCAAGACAGATCTATGACGATGGCACTTCTGG 211
      |||
Db      1  TCAACTAGCTGTTTATGTTAAGCCCAAGACAGATCTATGACGATGGCACTTCTGG 60

RESULT 10
US-10-059-585-5
; Sequence 5, Application US/10059585
; Publication No. US20030082776A1
; GENERAL INFORMATION:
; APPLICANT: Ota, Toshio
; APPLICANT: Isogai, Takao
; APPLICANT: Nishikawa, Tetsuo
; APPLICANT: Hayashi, Koji
; APPLICANT: Otsuka, Kaoru

```

```

; APPLICANT: Yamamoto, Jun-ichi
; APPLICANT: Ishii, Shizuko
; APPLICANT: Sugiyama, Tomoyasu
; APPLICANT: Makamatsu, Ai
; APPLICANT: Nagai, Keiichi
; APPLICANT: Otsuki, Tetsuji
; APPLICANT: Funahashi, Shin-ichi
; APPLICANT: Senoo, Chiaki
; APPLICANT: Nezu, Jun-ichi
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN
; TITLE OF INVENTION: KINASE/PROTEIN PHOSPHATASE
; FILE REFERENCE: 06501-098001
; CURRENT APPLICATION NUMBER: US/10/059,585
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05060
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/183,322
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: US 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 11-248036
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (173)...(1450)
; US-10-059-585-5

```

Query Match 3.0%; Score 42; DB 15; Length 2200;

Best Local Similarity 45.5%; Pred. No. 0.04; Mismatches 180; Indels 0; Gaps 0;

```

Matches 150; Conservative 0; Mismatches 180; Indels 0; Gaps 0;
QY 835 AATATATCCAGAACGTTGTAAGTGTGATGAGATCTGCACAGCTGGATGTTGGCAG 894
DB 158 AAGGATGAGAGGAGGATGTCAGAGGAGAGAGCTGCTGATGTTGATGAGGGCTGAGAG 217
QY 895 GGTGCTCATGATGATGCGCGTGGAGCTTTATATCATGAGGAACTCTCACTTATAA 954
DB 218 GCAGCCCTGACCCAGCCGAGGTGGGCGGAGAGGGGGTGGCTGCGCTCTTCAAGG 277
QY 955 GATCTGGAGTGCCTGCCAAGAGAGCTCTGCGGCTGGTCTGGAAGTGCAGAGAACAA 1014
DB 278 GAGAGAGGAGGAGCTCTGATGCAAGAGTGTGCTGCTGCTTCAAGTGAAGAGAGGA 337
QY 1015 GGTGAGTGTGCTCTTCCATATATATCAAGTAAATATTTCCACTACAGT 1074
DB 338 GAGAGAGAGAGTCCCTGAGATTTCTTGATACCAACCAAGGCTCTCGGCCCGACTAGC 397
QY 1075 CTGATGATGAGTGTGACGCGAAGCTTTTACCCACTGGGCTGCAATTACTGGCAGT 1134
DB 398 ATCCCTGCTCTTATATACAGACGTCGAGAACCAACGACCTGAGATGCTGACCGG 457
QY 1135 GAAGAAGCGAGGCGCATCATGAGAGAGTT 1164
DB 458 GAGAACACGTTTGTGTTAAGGTGAGAGT 487

```

# RESULT 11

```

US-09-908-975-1202
; Sequence 1202, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon

```

```

; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Irit
; APPLICANT: FAIGLER, Simon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLIC
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1202
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-09-908-975-1202

```

Query Match 2.9%; Score 41; DB 13; Length 65;

Best Local Similarity 76.9%; Pred. No. 0.0087; Mismatches 15; Indels 0; Gaps 0;

```

Matches 50; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 343 AGGGGAGAGAAATCAGCTGTGATGCTGAGACCAAGATTGATAGATGATGCTGGT 402
DB 1 AGGGGCAAGAAATCTGCAGTATGCTGCTGCTGCAATTCACAAGTTGCTATTAGGC 60
QY 403 CTTGG 407
DB 61 CTTGG 65

```

# RESULT 12

```

US-09-918-995-30889
; Sequence 30889, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30889
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(484)
; OTHER INFORMATION: n = A,T,C or G
; US-09-918-995-30889

```

Query Match 2.8%; Score 39.4; DB 11; Length 484;

Best Local Similarity 49.3%; Pred. No. 0.11; Mismatches 106; Indels 0; Gaps 0;

```

Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
QY 216 TACTGTGAGCCAGATGAGTGTCTCAAGAACTTGAAAAAGCCATCCAAATTATGTA 275
DB 231 TAGTGTCTGTTCCAAAGCACTGTGTACAGAACTGAAATTACATTAGAGGCAAC 290
QY 276 CCAAAACCTGAGCAAGATGAGGCTGAGAAAGTTCACTGAGGCACTGAGAAATACCCA 335
DB 291 CCTGAGAAATACACAGGGGCAAGCTTCCAGTATGATGTGGGAGAGAGAGGGCAG 350
QY 336 CTGGAGAGGGGAGAAATCAGCTGTGATGCTGAGCCCAAGATTATAGATAGCCAT 395
DB 351 AGGGGACAGGGGACAGATTCAGCTTGTGTGGGTCTGAGGGGTCTACAGGGGGTAG 410

```

```
Oy      396 CCTGGGCTTTGGCAGCAGCATTTGGACTC   424  
          ||| | | | |  
Db     411 CCAGATCTGGGAACAGATCAGCGACTC   439
```

RESULT 13  
US-10-232-484-5

; GENERAL INFORMATION:

```

; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: No. US20030113847A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 791CIP2B1V
; CURRENT APPLICATION NUMBER: US/10/232,484

```

OTHER INFORMATION:  
US-10-232-484-5

	Query Match	2.8%	Score 39.4	DB 15	Length 1363
	Best Local Similarity	49.3%	Pred No. 0.2	Mismatches 106	Indels 0
	Matches 103	Conservative	0	Matches 106	Gaps 0
QY	216	TACTGTTGACCCAGACTGAGTGGCTCCAGAACCTAGAAAAAGCATCCAAATTATGTA	275		
Db	877	TAGTGTCTGTTCCAAACCACTGTGTCAACAGAGCTAGAAATTACATTAGAGGCACAAAC	936		
QY	276	CCAAACCTGCAGCAAGATGGGCTGTGAGAAATTCACTGCAGCCAGTGAATACCCCA	335		
Db	937	CCCTGAGAAATACAAAGGGGGGACCTTCCATTAAGATGTGTTGGGAAAGAGAGGGCAG	996		
QY	336	CTGGAGAGGGGAGAAAGATCACTGTGTGCTGTGAGCCAGAAATTCATTAAGATAGCAT	395		
Db	997	AGGGGACAGGGGACAGGATTCAGCTTTTGTGTGTGGATCCGTAGAGGTTCTCTACAGGGGTAG	1055		
QY	396	CCTGGGCTTTGGCAGCAGCATTTGGGACATC	424		
Db	1057	CCAGGATCTGGGAAACAGATCAGCACCATC	1085		

RESULT 14  
US-10-044-090-579

GENERAL INFORMATION:

```

1  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
2
3  FILE REFERENCE: PA-0028 US
4
5  CURRENT APPLICATION NUMBER: US/10/044,090
6
7  CURRENT FILING DATE: 2002-01-09
8
9  NUMBER OF SEQ ID NOS: 850
10

```

```

; OTHER INFORMATION: Incyte ID No. US20020137081A1 023344.1.
; NAME/KEY: unsure
; LOCATION: 169, 173
; OTHER INFORMATION: a, b, c, g, or other
US-10-044-090-579

```

Query Match	2.8%;	Score 39.4;	DB 14;	Length 2319;
Best Local Similarity	49.3%;	Pred. No. 0.29;		
Matches 103; Conservative	0;	Mismatches 106;	Indels 0;	Gaps 0;

QY	216	TACGTGTGGACCCGAGACTGAGTGGCTCCCAAAACCTAGAAAAAGCATCCAAATATGTA	275
Db	1560	TAGTGTCTGTTCCAAAGCACTGTGTACAGAGCTAGAAATTACATTAGAGCCAAAC	1611
QY	276	CCAAACCTCGACGACAGATGGGCTGAGAAAATTCACTCGAGCCAGTGAATATACCCA	335
Db	1620	CCCTGAGAAATACACAAAGGGGCAAGCTTCCAGTATGATGTGTTGGGGAAGAGAGGAGCAG	1675
QY	336	CTGGGAGAGGGGAGAGATCAGCTGTGATCTCTGAGCCAGAAATTCTAAGATAGCCAT	395
Db	1680	AGGGGACACGGGGACAGAGATTCAGCTGTGTGGTGCGTCTCTGAGGGTTCCTACACAGGGGTAG	1735
QY	396	CCGTGGCTTTGGCAGCAGCAATTGGGACATC	424
Db	1740	CCAGGATCTGGGAAACAGATCAGCAGATC	1768

RESULT 15  
US-09-814-353-20277  
; Sequence 20277, Application US/09814353  
; Publication No. US20030165831A1

APPLICANT: Lee, John  
; APPLICANT: Thompson, Pamela

TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER

```

1 CURRENT APPLICATION NUMBER: US-09/814,353
2 CURRENT FILING DATE: 2001-03-21
3 PRIOR APPLICATION NUMBER: US 60/191,031
4 PRIOR FILING DATE: 2000-03-21
5 PRIOR APPLICATION NUMBER: US 60/207,124
6 PRIOR FILING DATE: 2000-03-25
7 PRIOR APPLICATION NUMBER: US 60/211,940
8 PRIOR FILING DATE: 2000-06-15
9 PRIOR APPLICATION NUMBER: US 60/216,820
10 PRIOR FILING DATE: 2000-07-07
11 PRIOR APPLICATION NUMBER: US 60/220,661
12 PRIOR FILING DATE: 2000-07-25
13 PRIOR APPLICATION NUMBER: US 60/257,672
14 PRIOR FILING DATE: 2000-12-21
15 NUMBER OF SEQ ID NOS: 22037
16 SOFTWARE: FastSeq for Windows Version 4.0
17 SEQ ID NO 20277
18
19 LENGTH: 3332
20
21 TYPE: DNA
22
23 ORGANISM: Homo sapiens
24
25 US-09-814-353-20277

```

Query Match	2.8%	Score 39.4;	DB 13;	Length 3322;
Best Local Similarity	49.3%	Pred. No. 0.36;		
Matches 103; Conservative	0;	Mismatches 106;	Indels 0;	Gaps 0

Qy	216	TACGTTGGAACCGACGCTGAGTGGCTCCCAACACCTGAGAAAAAGCATTCCAAATATATGTA	275
Db	1946	TAGGTCTCTGTTCCAAAGCAGCTGTGTACAGAGCTAGAAATTACATTAGAGCGCAAAAC	2009
Qy	276	CCAAACCTCGACAGATGGGCTGGAGAAAGTTTCACCTGAGCCAGTGAAGATATCCCA	335
Db	2006	CCCTGAGATATCAACAGGGGGCAGCTTCCAGTATGTGTTGGGAGAGAGAGAGCGCAG	206

Qy	336	CTGGGAGAGGGGAGAGATCAGCTGTGATGCTGGAGCCAGAAATTCATTAAGATAGCCAT	395
Db	2066	AGGGGACAGGGGACAGGATTCAGCTTGTGTGGGTCTGAGGGTTCCTAACAGGGGTAG	2125
Qy	396	CCTGGGTCTTGGGACAGCATTTGGGACTC	424
Db	2126	CCAGGATCTGGGAAACAGATCAGCGACTC	2154

Search completed: December 22, 2003, 23:44:40  
Job time : 447.263 secs

**THIS PAGE BLANK (USPTO)**



Tue Dec 23 11:24:30 2003

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

.OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:18:08 ; Search time 4851.69 Seconds  
(without alignments)  
11332.658 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_171\_1514  
Perfect score: 1344  
Sequence: 1 AAGAAATGCATCTCTAAGAG.....AAGAAATGCTGCTAGGTCC 1344

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GeneBml:.\*  
1: gb Da:.\*  
2: gb Htg:.\*  
3: gb In:.\*  
4: gb Om:.\*  
5: gb Ov:.\*  
6: gb Pat:.\*  
7: gb Ph:.\*  
8: gb Pl:.\*  
9: gb Pr:.\*  
10: gb Ro:.\*  
11: gb Sts:.\*  
12: gb Sy:.\*  
13: gb Un:.\*  
14: gb Vi:.\*  
15: em Ba:.\*  
16: em Fun:.\*  
17: em Hum:.\*  
18: em In:.\*  
19: em Mu:.\*  
20: em Om:.\*  
21: em Or:.\*  
22: em Ov:.\*  
23: em Pat:.\*  
24: em Ph:.\*  
25: em Pl:.\*  
26: em Ro:.\*  
27: em Sts:.\*  
28: em Un:.\*  
29: em Vi:.\*  
30: em Htg\_Hum:.\*  
31: em Htg\_Inv:.\*  
32: em Htg\_Other:.\*  
33: em Htg\_Mus:.\*  
34: em Htg\_Pln:.\*  
35: em Htg\_Rod:.\*  
36: em Htg\_Mam:.\*  
37: em Htg\_Vrt:.\*  
38: em Sv:.\*  
39: em Htgo\_Hum:.\*  
40: em Htgo\_Mus:.\*  
41: em Htgo\_Other:.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	100.0	1851	6	BD106411 Secreted
2	1340.8	99.8	1778	6	BD105816 Secretory
3	1340.8	99.8	1860	6	AX006440 Sequence
4	1340.8	99.8	1928	9	BC020689 Homo sapi
5	1337.6	99.5	1472	9	AF107834 Homo sapi
6	1336	99.4	1860	6	BD127520 Primer fo
7	1336	99.4	1860	6	AK075132 Homo sapi
8	1327.2	98.8	1794	9	AF119386 Homo sapi
9	1086.4	80.7	1596	10	BC037067 Mus muscu
10	1084.8	80.7	1596	10	AF107835 Mus muscu
11	1060.8	78.9	1726	10	AF131077 Rattus no
12	1060.8	78.9	1778	6	AX400829 Sequence
13	1060.8	78.9	1778	10	AF097723 Rattus no
14	1000.4	74.4	1716	10	AF009513 Mus muscu
15	448.6	33.4	609	6	BD059610 Secreted
16	421.4	31.4	423	9	AF107833 Homo sapi
17	398.2	29.6	642	6	BD125210 Primer fo
18	398.2	29.6	642	6	BD126476 Homo sapi
19	364.4	27.1	156702	2	AC084223 Homo sapi
20	364.4	27.1	157927	9	AP006278 Homo sapi
21	364.4	27.1	171301	9	AC010859 Homo sapi
22	331	24.6	417	10	AF107832 Mus muscu
23	292.8	21.8	214287	2	AC129951 Mus muscu
24	279.4	20.8	167883	2	AC121026 Rattus no
25	279.4	20.8	218874	2	AC110103 Rattus no
26	212.8	15.8	314	6	BD058492 Secreted
27	212.2	15.8	144703	9	AP003112 Homo sapi
28	205	15.3	152176	9	AP003117 Homo sapi
29	205	15.3	171035	9	AP003111 Homo sapi
30	204.6	15.2	462	6	BD105811 Secretory
31	204.6	15.2	462	6	E62849 Secretory p
32	174.4	13.0	214287	2	AC129951 Mus muscu
33	172.8	12.9	203469	2	AC123654 Mus muscu
34	171.2	12.7	218874	2	AC110103 Rattus no
35	163.4	12.2	166050	2	AC013817 Homo sapi
36	163.4	12.2	177008	2	AC011175 Homo sapi
37	163.4	12.2	240178	2	AC094173 Rattus no
38	163.4	12.2	288136	2	AC120489 Rattus no
39	154.4	11.5	238034	2	AC123681 Mus muscu
40	140.6	10.5	317	6	BD076934 5' EST of
41	116	8.6	10792	1	AE015790 Shewanell
42	115	8.6	110000	2	AL390202_01 Continuation (2 of
43	114.2	8.5	492	6	BD125961 Primer fo
44	99.2	7.4	10487	1	AE005922 Caulobact
45	96.4	7.2	1729	3	AF077194 Acanthoch

## ALIGNMENTS

RESULT 1  
LOCUS BD106411 1851 bp DNA linear PAT 18-SEP-2002  
DEFINITION Secreted proteins and polynucleotides encoding them.  
ACCESSION BD106411  
VERSION BD106411.1 GI:23201229  
KEYWORDS JP 2002503955-A/2.  
SOURCE Chlamydia sp.  
ORGANISM Chlamydia sp.  
REFERENCE 1 (bases 1 to 1851)  
JACOBS,K., MCCOY,J.M., LAVALLIE,E.R., RACIE,L.A., MERBERG,D.,  
TREACY,M., SPAULDING,V. and AGOSTINO,M.J.  
Secreted proteins and polynucleotides encoding them  
Patent: JP 2002503955-A 2 05-FEB-2002;  
JOURNAL

## COMMENT

GENETICS INSTITUTE INC  
PN JP 2002503955-A/2  
PD 05-FEB-2002  
PF 20-MAR-1998 JP 1998545874  
PR 21-MAR-1997 US 08/822167,19-MAR-1998 US 09/044466 PI  
KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE, PI  
DAVID MERBERG,  
PI MAURICE TREACY,VIKKI SPAULDING,MICHAEL J AGOSTINO PC  
CI2N15/12,C07K14/47,A61K38/17  
CC Strandedness: Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.

## FEATURES

source  
1..1851  
/organism="Chlamydia sp."  
/mol\_type="genomic DNA"  
/db\_xref="taxon:35827"  
BASE COUNT 531 a 413 c 438 g 469 t  
ORIGIN

Query Match 100.0%; Score 1344; DB 6; Length 1851;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGAATGCGATCTCTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGA 60  
DB 171 AAGAATGCGATCTCTAAGAGGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGA 230  
QY 61 GATGTGCTAAGCAATCACTAGCTGTTATAGTAAAGCCAGAACAGATCTAT 120  
DB 231 GATGTGCTAAGCAATCACTAGCTGTTATAGTAAAGCCAGAACAGATCTAT 290  
QY 121 GAGGATTTGGAAGCTTCTGTTGATCTGTGGAAGCCAGAGTGGCTCCAGAACTTA 180  
DB 291 GAGGATTTGGAAGCTTCTGTTGATCTGTGGAAGCCAGAGTGGCTCCAGAACTTA 350  
QY 181 GAAAAAGCCATCAATATATGTAACCAAAAGCTGAGCAAGATGGCTGGAAGATTGAC 240  
DB 351 GAAAAAGCCATCAATATATGTAACCAAAAGCTGAGCAAGATGGCTGGAAGATTGAC 410  
QY 241 CTGAGGCGAGTGAATAACCCCACTGGAGAGGAGAGAAATCAAGTGTGTGAG 300  
DB 411 CTGAGGCGAGTGAATAACCCCACTGGAGAGGAGAGAAATCAAGTGTGTGAG 470  
QY 301 CCAAGAATTCATATAGATAGCCATCTGCTGCTTGGCAGCAGATGGGATCTCTCAAA 360  
DB 471 CCAAGAATTCATATAGATAGCCATCTGCTGCTTGGCAGCAGATGGGATCTCTCAAA 530  
QY 361 GGCATTACAGCAGAAAGTTCTGGTGTGAAGCTTCTGATGAAGCTGCAAGAGGCTCA 420  
DB 531 GGCATTACAGCAGAAAGTTCTGGTGTGAAGCTTCTGATGAAGCTGCAAGAGGCTCA 590  
QY 421 GAAGCAAGAGGAGATGTTGTTTATACCAACCTTACATCACTACTCAAGAGCGGTG 480  
DB 591 GAAGCAAGAGGAGATGTTGTTTATACCAACCTTACATCACTACTCAAGAGCGGTG 650  
QY 481 CAATATCCCAAGCGAGGCGGTGTGAAGCTGCAAGAGTGGGCTTTGGCATCTCTCAT 540  
DB 651 CAATATCCCAAGCGAGGCGGTGTGAAGCTGCAAGAGTGGGCTTTGGCATCTCTCAT 710  
QY 541 CGATCCGTGGCTCTCTTCTCATCTACAGTCTCTCACAGGATTCAGAAATACAGGAT 600  
DB 711 CGATCCGTGGCTCTCTTCTCATCTACAGTCTCTCACAGGATTCAGAAATACAGGAT 770  
QY 601 GGCCTGCCCAAAATTCACAGCCTGTATTTACGTGGAAGATGCAAGAAATGATCAAGA 660  
DB 771 GGCCTGCCCAAAATTCACAGCCTGTATTTACGTGGAAGATGCAAGAAATGATCAAGA 830  
QY 661 ATGGCTTCTCATGGATCAAAATTTGTCTATTCAGCTAAAGATGGGGCAAAAGCTTACCA 720  
DB 831 ATGGCTTCTCATGGATCAAAATTTGTCTATTCAGCTAAAGATGGGGCAAAAGCTTACCA 890  
QY 721 GATACTGATTCCTTCAACACTGTACAGATCACTGGGAGCAATATTCAGAAACAGTT 780

DB 891 GATACTGATTCCTTCAACACTGTACAGATCACTGGGAGCAAAATATCCAGAACAGTT 950  
QY 781 GTACTGTGCTAGTGAACATCTGAGACAGTGGGATTTGGGCGAGGCTCCATGATGATGC 840  
DB 951 GTACTGTGCTAGTGAACATCTGAGACAGTGGGATTTGGGCGAGGCTCCATGATGATGC 1010  
QY 841 GGTGAGCCCTTTATATATATATGAGGAAAGCACTCACTTATTAAGATCTTGGGCTGCTCA 900  
DB 1011 GGTGAGCCCTTTATATATATATGAGGAAAGCACTCACTTATTAAGATCTTGGGCTGCTCA 1070  
QY 901 AAGAGACTGCGGCTGTGCTCTGAGACTGCAAGAGAACAGGTGAGTTGCTCTTC 960  
DB 1071 AAGAGACTGCGGCTGTGCTCTGAGACTGCAAGAGAACAGGTGAGTTGCTCTTC 1130  
QY 961 CAGTATATCAGTTACCAAGATTAATTTTCCACTACAGTCTGTGATGAGTCTGAC 1020  
DB 1131 CAGTATATCAGTTACCAAGATTAATTTTCCACTACAGTCTGTGATGAGTCTGAC 1190  
QY 1021 GCAGGAACCTTCTTACCCACTGGGCTGCAATTCATCTGCGAGTGAAGGCGAGGCGCATC 1080  
DB 1191 GCAGGAACCTTCTTACCCACTGGGCTGCAATTCATCTGCGAGTGAAGGCGAGGCGCATC 1250  
QY 1081 ATGAGAGAGGTTATGAGCTGTGCAAGCCCTCAATATCACTCAGTCTGAGCCATGGA 1140  
DB 1251 ATGAGAGAGGTTATGAGCTGTGCAAGCCCTCAATATCACTCAGTCTGAGCCATGGA 1310  
QY 1141 GAAGGACAGACATCAACTTTTGGATCCAAAGCTGAGATGCTTGGAGCCATCTTATGAT 1200  
DB 1311 GAAGGACAGACATCAACTTTTGGATCCAAAGCTGAGATGCTTGGAGCCATCTTATGAT 1370  
QY 1201 GACTTATACAGATATTTCTTCTTCATCACTCCACGAGAGCAACATGATCATGAT 1260  
DB 1371 GACTTATACAGATATTTCTTCTTCATCACTCCACGAGAGCAACATGATCATGAT 1430  
QY 1261 CCAAGCAGATGATGTTGCTGTGCTGTGTTGGCTGTGTTCTTATGTTGTGACAGC 1320  
DB 1431 CCAAGCAGATGATGTTGCTGTGCTGTGTTGGCTGTGTTCTTATGTTGTGACAGC 1490  
QY 1321 ATGGAAGAAATGCTGCTAGGTCC 1344  
DB 1491 ATGGAAGAAATGCTGCTAGGTCC 1514

## RESULT 2

BD105816

LOCUS BD105816 1778 bp DNA linear PAT 18-SEP-2002

DEFINITION Secretory protein.

ACCESSION BD105816  
VERSION BD105816.1 GI:23200634

KEYWORDS UP 2002502234-A/7.

SOURCE

ORGANISM  
synthetic construct  
artificial sequences.

REFERENCE 1 (bases 1 to 1778)

AUTHORS Jacobs K., McCoy, J.M., Racie, L.A., Lavallie, E.R., Merberg, D. and

Spaulding, V.

TITLE Secretory protein

JOURNAL Patent: JP 2002502234-A 7 22-JAN-2002;

## COMMENT

GENETICS INSTITUTE INC

PN JP 2002502234-A/7

PD 22-JAN-2002

PF 16-APR-1997 JP 1997537384

PR 18-APR-1996 US 08/634325,13-JAN-1997 US 08/783520 PI

KENNETH JACOBS,JOHN M MCCOY,LISA A RACIE,EDWARD R LAVALLIE, PI

DAVID MERBERG,

PI VIKKI SPAULDING

PC CI2N15/12,C07K14/47,A61K38/17

CC Strandedness: Double;

CC Topology: Linear;

FH Key Location/Qualifiers.

FEATURES

source 1..1778

/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
BASE COUNT 514 a 386 c 422 g 456 t  
ORIGIN

Query Match 99.8%; Score 1340.8; DB 6; Length 1778;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 AAGAAATGGCAATCTCTAAGAGACCTTTGAAAGAAATTAAGAAATAGCCAGCTGGA 60
DB AAGAAATGGCAATCTCTAAGAGACCTTTGAAAGAAATTAAGAAATAGCCAGCTGGA 148
QY 61 GATGTTCTAAGCAATCATCAACCTAGCTGTTATGATGATTAAGCCAGAAACAGATCTAT 120
DB 149 GATGTTCTAAGCAATCATCAACCTAGCTGTTATGATGATTAAGCCAGAAACAGATCTAT 208
QY 121 GAGCGATTGGCACTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 180
DB 209 GAGCGATTGGCACTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 268
QY 181 GAAAAAGCCATCCAAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 269 GAAAAAGCCATCCAAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 328
QY 241 CTGAGGCAATGAGAAATACCCCACTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 329 CTGAGGCAATGAGAAATACCCCACTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 388
QY 301 CCAAGAAATTAAGATAGGCAATCTGAGTCTTGGCAGAGAGATGGGAGATCTCCAGAA 360
DB 389 CCAAGAAATTAAGATAGGCAATCTGAGTCTTGGCAGAGAGATGGGAGATCTCCAGAA 448
QY 361 GGCATTAACAGCAAGATGTTCTGTTGATGATGATGATGATGATGATGATGATGATGAT 420
DB 449 GGCATTAACAGCAAGATGTTCTGTTGATGATGATGATGATGATGATGATGATGATGAT 508
QY 421 GAAGCAAGAGAGAGATGTTGTTTATTAACCACTTAACCTTAACCTTAACCTTAACCT 480
DB 509 GAAGCAAGAGAGAGATGTTGTTTATTAACCACTTAACCTTAACCTTAACCTTAACCT 568
QY 481 CAATACCGAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 569 CAATACCGAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 628
QY 541 CGATCCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
DB 629 CGATCCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 688
QY 601 GGGCGTCCCAAAATTTCCAAAGCTGTTATTAACCGTGGAGAGATGAGAAATGATGAT 660
DB 689 GGGCGTCCCAAAATTTCCAAAGCTGTTATTAACCGTGGAGAGATGAGAAATGATGAT 748
QY 661 ATGGCTTCTATGGGATCAAAATTTGATGATGATGATGATGATGATGATGATGATGAT 720
DB 749 ATGGCTTCTATGGGATCAAAATTTGATGATGATGATGATGATGATGATGATGATGAT 808
QY 721 GATACGATTTCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTTCCAGAAACAG 780
DB 809 GATACGATTTCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTTCCAGAAACAG 868
QY 781 GATACGATTTCTTCAACACTGTAGCAGAGATGATGATGATGATGATGATGATGATGAT 840
DB 869 GATACGATTTCTTCAACACTGTAGCAGAGATGATGATGATGATGATGATGATGATGAT 928
QY 841 GGTGAGAGCTTTTATATCAATGAGAGACTCTCACTTATTAAGATCTTGGCTGCTGCA 900
DB 929 GGTGAGAGCTTTTATATCAATGAGAGACTCTCACTTATTAAGATCTTGGCTGCTGCA 988
QY 901 AAGAGAGCTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 989 AAGAGAGCTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1048

```

```

QY 961 CAGTATTATCAGTTATACACAGATTAATATTTCACATACAGTCTGTATGATGATGATGAT 1020
DB 1049 CAGTATTATCAGTTATACACAGATTAATATTTCACATACAGTCTGTATGATGATGATGAT 1108
QY 1021 GCAGAACTTTCTTACCCTGAGCTGAGATTAATCACTGAGAGAGAGAGAGAGAGAGAG 1080
DB 1109 GCAGAACTTTCTTACCCTGAGCTGAGATTAATCACTGAGAGAGAGAGAGAGAGAGAG 1168
QY 1081 ATGAGAGAGTTATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1140
DB 1169 ATGAGAGAGTTATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1228
QY 1141 GAAGGAGACATCACTCACTTTGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1229 GAAGGAGACATCACTCACTTTGATGATGATGATGATGATGATGATGATGATGATGAT 1288
QY 1201 GACTTATACAAATTTCTTCTTCACTGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 1289 GACTTATACAAATTTCTTCTTCACTGATGATGATGATGATGATGATGATGATGATGAT 1348
QY 1261 CCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1349 CCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1408
QY 1321 ATGAGAGAAATGCTGCTAGGTC 1344
DB 1409 ATGAGAGAAATGCTGCTAGGTC 1432

RESULT 3
AX006440
LOCUS
DEFINITION
Sequence 5 from Patent WO0004157.
AX006440
VERSION
AX006440.1 GI:9994575
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
Neeffs,J.M., Peeters,D.C. and Pangalos,M.
Cloning and characterisation of novel mammalian peptidases
JOURNAL
PATENT: WO 0004157-A 5 27-JAN-2000;
JANSSEN PHARMACEUTICA NV (BE); NEEFS JEAN MARC EDMOND PERNAND (BE);
PEETERS DANIELLE CELINE GEORGE (BE); PANGALOS MENLAS (GB)
FEATURES
Location/Qualifiers
1..1860
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 518 a 420 c 451 g 471 t
ORIGIN
Query Match 99.8%; Score 1340.8; DB 6; Length 1860;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Db      377  GAAAAAGCCATCAAAATTATGATCCAAAACCTGCAGCAAGATGGGCTGGAGAAAGTTCA 436
Qy      241  CTGAGCCAGTGAAGTAATCCCACTGGAGAGAGGAGAAAGTAATCAAGCTGTGATGTGGAG 300
Db      437  CTGGAGCGAGAGAAATACCCCACTGGAGAGAGGAGAAAGTAATCAAGCTGTGATGTGGAG 496
Qy      301  CCAAGAAATTCATTAAGATAGCCATCTGGGCTTGGCAGACAGATGGGAGCTCTCCAGAA 360
Db      497  CCAAGAAATTCATTAAGATAGCCATCTGGGCTTGGCAGACAGATGGGAGCTCTCCAGAA 556
Qy      361  GGCATTACAGCAAGAGTTCTGGTGTGA 377TGATGAATCTGCAGAGAAAGGCTCA 420
Db      557  GGCATTACAGCAAGAGTTCTGGTGTGA 377TGATGAATCTGCAGAGAAAGGCTCA 616
Qy      421  GAAGCAAGAGGAAAGTTGTTTATATACCACTTACATCACTACTCAAGAGCGGTG 480
Db      617  GAAGCAAGAGGAAAGTTGTTTATATACCACTTACATCACTACTCAAGAGCGGTG 676
Qy      481  CAATACCGAAGCGAGGGGGGGTGGAGCTGCAGAGTGGGGGCTTTGGCATCTCTCAT 540
Db      677  CAATACCGAAGCGAGGGGGGGTGGAGCTGCAGAGTGGGGGCTTTGGCATCTCTCAT 736
Qy      541  CGATCCGTCGCTCTCTCTCTCAATCAAGTCTCA 377TGATGAATCTGCAGAGAA 600
Db      737  CGATCCGTCGCTCTCTCTCTCTCAAGTCTCA 377TGATGAATCTGCAGAGAA 796
Qy      601  GGGCGTCCCAAAATTCACAGAGCTGTATACGCTGTGAAGATGCAGAAATGATGCAAGA 660
Db      797  GGGCGTCCCAAAATTCACAGAGCTGTATACGCTGTGAAGATGCAGAAATGATGCAAGA 856
Qy      661  ATGGCTTCATGAGGATCAAAATTTGATCAATAGCTAAAGATGGGGGGCAAGACCTACCA 720
Db      857  ATGGCTTCATGAGGATCAAAATTTGATCAATAGCTAAAGATGGGGGGCAAGACCTACCA 916
Qy      721  GATACCTGATCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAAAGGTT 780
Db      917  GATACCTGATCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAAAGGTT 976
Qy      781  GTACTGGTCAATGCAATCTGCAAGCTGGAGATTTGGGCGAGGTGCAATGATGATGCG 840
Db      977  GTACTGGTCAATGCAATCTGCAAGCTGGAGATTTGGGCGAGGTGCAATGATGATGCG 1036
Qy      841  GGTGGAGCTTATATCAATGGAAGACACTCACTTATTAAGATCTTGGGCTGGGTGCA 900
Db      1037  GGTGGAGCTTATATCAATGGAAGACACTCACTTATTAAGATCTTGGGCTGGGTGCA 1096
Qy      901  AAGAGACTCTGCGGCTGTGCTGTGA 377TGATGAAGAAAGAGTGGATGTGCTTC 960
Db      1097  AAGAGACTCTGCGGCTGTGCTGTGATGAGAGAAAGAGAGTGGATGTGCTTC 1156
Qy      961  CAGTATTCAGTTTACAGAGTAAATTTTCCAACTCAAGTCTGTGATGAGAGTCTGAC 1020
Db      1157  CAGTATTCAGTTTACAGAGTAAATTTTCCAACTCAAGTCTGTGATGAGAGTCTGAC 1216
Qy      1021  GCAGAGACCTTCTTAACCACTGGGCTGCAATTCATGAGAGTAAAGAGGAGGAGCTATC 1080
Db      1217  GCAGAGACCTTCTTAACCACTGGGCTGCAATTCATGAGAGTAAAGAGGAGGAGCTATC 1276
Qy      1081  ATGAGAGAGTATGAGCTGTGAGAGGCTCTCAATATCACTCAAGTCTGAGCCATGGA 1140
Db      1277  ATGAGAGAGTATGAGCTGTGAGAGGCTCTCAATATCACTCAAGTCTGAGCCATGGA 1336
Qy      1141  GAAGGAGACAGATCACTTTTGGATCCAAGCTGAGAGTCTGAGAGCAAGTCTTACTTGAT 1200
Db      1337  GAAGGAGACAGATCACTTTTGGATCCAAGCTGAGAGTCTGAGAGCAAGTCTTACTTGAT 1396
Qy      1201  GACTTATCAAGTATTTCTCTTCATCACTCCCAAGGAGACACCATGATGATGATGAT 1260
Db      1397  GACTTATCAAGTATTTCTCTTCATCACTCCCAAGGAGACACCATGATGATGATGAT 1456
Qy      1261  CCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Db      1457  CCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1516

```

```

Qy      1321  ATGGAAGAAATGCTGCTAGGTCC 1344
Db      1517  ATGGAAGAAATGCTGCTAGGTCC 1540

RESULT 4
BC020689
LOCUS      BC020689
DEFINITION Homo sapiens, plasma glutamate carboxypeptidase, clone MGC:22418
IMAGE:4251802, mRNA, complete cds.
ACCESSION BC020689
VERSION    BC020689.1 GI:18088383
KEYWORDS   MGC.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 1928)
            Strausberg, R.
            Direct Submission
            Submitted (03-JUN-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
            NIH-MGC Project URL: http://mgc.nci.nih.gov
            Contact: MGC help desk
            Email: cgabs-remail.nih.gov
            Tissue Procurement: ATCC/DCT/DRP
            CDNA Library Preparation: CLONTECH Laboratories, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNU)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA 94305
            Web Site: http://www.shgc.stanford.edu
            Contact: (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
            R. M.

FEATURES
     source
     location/Qualifiers
         1..1928
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="LocusID:10404"
            /db_xref="taxon:9606"
            /clone="MGC:22418 IMAGE:4251802"
            /tissue_type="Skin, melanotic melanoma, high MDR."
            /clone_id="NH_MGC_62"
            /lab_host="DH10B"
            /note="Vector: pDNR-LIB"
            157..1585
            /codon_start=1
            /product="plasma glutamate carboxypeptidase"
            /protein_id="AAH20689.1"
            /db_xref="GI:18088384"
            /translation="MKFLIPAFPGVHLISLGSKAICKNGISKRFPFEIKETIASCG
            DVAKIINILAVGKAORNSYERLALIVDTGPLPSKNEKAIQIMYOLQODGLEK
            VHEPRVPIHMERGESBAVWLBERIKHIALIGGSSIGTPPEGITAEVLVYTSDEIQ
            RRASEARGKIVNQPIYNSRTVQRTQGAVEAAKAGALASLRVSFISYBHTG
            IREYODVPRKIPACTIVEDAEMMSRMSAGIKIVIQKXGATYDPTDSFNVAEIT
            GSKYPEGVIVVSGHLDSDMGQAMDGGAFISWEALSLIKDLGAPKRTLRVLEMT
            AEROGVAGFOYVQLHKVINSYSLVMSAGTFLPGLOFTSEKRAIMEEYMSLL
            OPNTQVYSHGSGTDINFWIOAGVPRASLLDLKYKFFHHSHGDMTMDPKQNV
            AAAYVAVSYVAVDMEMLPRS"

BASE COUNT      554 a      435 c      466 g      473 t
ORIGIN
Query Match      99.8%; Score 1340.8; DB 9; Length 1928;
Best Local Similarity 99.9%; Pred. No. 0;

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNU at: <http://image.llnl.gov>  
 Series: IRAL Plate: 35 Row: f Column: 13  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7706386.

Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 AAGATGCACTCTTAAGAGACTTTTGAAGAATAAAGAAATAGCAGCTGGGA 60
Db 239 AAGATGCACTCTTAAGAGACTTTTGAAGAATAAAGAAATAGCAGCTGGGA 298
QY 61 GATGTTCTAAAGCAATCATCACTAGCTTTATGTTAAAGCCAGAAAGATCTAT 120
Db 299 GATGTTCTAAAGCAATCATCACTAGCTTTATGTTAAAGCCAGAAAGATCTAT 358
QY 121 GAGGATTTGGCACTTCTGTTGATTAATCTGTGAGACCCAGCTGATGGCTCCAGAACTTA 180
Db 359 GAGGATTTGGCACTTCTGTTGATTAATCTGTGAGACCCAGCTGATGGCTCCAGAACTTA 418
QY 181 GAAAAAGCCATCCAAATTTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAGTTTAC 240
Db 419 GAAAAAGCCATCCAAATTTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAGTTTAC 478
QY 241 CTGAGGCAAGTGAATATCCCACTGGAGAGAGGAGAGAAATCAGCTGTGATCTGGAG 300
Db 479 CTGAGGCAAGTGAATATCCCACTGGAGAGAGGAGAGAAATCAGCTGTGATCTGGAG 538
QY 301 CCAAGATTTCTAAGATAGCCATCTGGGCTTTGGCAGCAGATTTGGGATCTCTCCAGAA 360
Db 539 CCAAGATTTCTAAGATAGCCATCTGGGCTTTGGCAGCAGATTTGGGATCTCTCCAGAA 598
QY 361 GGCATTACAGAGAAAGTTCTGTGATGACCTTTTCATGATGAACTGAGAGAGGAGCTCA 420
Db 599 GGCATTACAGAGAAAGTTCTGTGATGACCTTTTCATGATGAACTGAGAGAGGAGCTCA 658
QY 421 GAAGCAAGAGAGAAAGTTGTTTATTAACCACTTACATCACTACTCAAGAGAGGAG 480
Db 659 GAAGCAAGAGAGAAAGTTGTTTATTAACCACTTACATCACTACTCAAGAGAGGAG 718
QY 481 CAATACCGAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
Db 719 CAATACCGAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 778
QY 541 CGATCGTGGGCTCTCTCATCTACAGTCTGACAGGATTTCCAGAAATACAGAGAT 600
Db 779 CGATCGTGGGCTCTCTCATCTACAGTCTGACAGGATTTCCAGAAATACAGAGAT 838
QY 601 GGGGTCGCCAAATTTCCAGAGCTGTATTAAGGTTGAGAGAGTGAAGATGATGTCAAGA 660
Db 839 GGGGTCGCCAAATTTCCAGAGCTGTATTAAGGTTGAGAGAGTGAAGATGATGTCAAGA 898
QY 661 ATGGCTTCTCATGGGATCAAAATTTGCTATTCAGTAAAGATGGGGCAAGACTTACCCA 720
Db 899 ATGGCTTCTCATGGGATCAAAATTTGCTATTCAGTAAAGATGGGGCAAGACTTACCCA 958
QY 721 GATACGATTCCTTCAACAGCTGTAGCAGAGATCTGAGGAGCAATATCCAGAAACAGATT 780
Db 959 GATACGATTCCTTCAACAGCTGTAGCAGAGATCTGAGGAGCAATATCCAGAAACAGATT 1018
QY 781 GTACTGTCAGTGAACATCTGAGCAGCTGGGATTTGGGCGAGGCTGCATGATGATGAGC 840
Db 1019 GTACTGTCAGTGAACATCTGAGCAGCTGGGATTTGGGCGAGGCTGCATGATGATGAGC 1078
QY 841 GGTGAGACCTTTATATCATGGGAAACACTCTCACTTATTAAGATCTTGGGCTGCTGCA 900
Db 1079 GGTGAGACCTTTATATCATGGGAAACACTCTCACTTATTAAGATCTTGGGCTGCTGCA 1138
QY 901 AAGAGCACTGCGGGCTGTGCTGTGACCTGCAAGAAAGAGTGAAGTTGTGCTTC 960
Db 1139 AAGAGCACTGCGGGCTGTGCTGTGACCTGCAAGAAAGAGTGAAGTTGTGCTTC 1198
QY 961 CAGTATTATGATTAACAAGATTAATTTCAACTTCAAGTCTGTGATGATGATGATGATGAT 1020
Db 1199 CAGTATTATGATTAACAAGATTAATTTCAACTTCAAGTCTGTGATGATGATGATGATGAT 1258
QY 1021 GCAGGAACCTTTTACCACTGGGCTGCAATTCATGCGAGTGAAGAAAGCCAGGCGCATC 1080
Db 1259 GCAGGAACCTTTTACCACTGGGCTGCAATTCATGCGAGTGAAGAAAGCCAGGCGCATC 1318

```

```

QY 1081 ATGAGAGAGGTTATGAGCTGTGACAGCCCTCAATATCACTGAGGCTGAGCATGGA 1140
Db 1319 ATGAGAGAGGTTATGAGCTGTGACAGCCCTCAATATCACTGAGGCTGAGCATGGA 1378
QY 1141 GAAAGGACAGACATCAACTTTTGATCCAGAGCTGAGAGCTGAGGAGCTTACTTAT 1200
Db 1379 GAAAGGACAGACATCAACTTTTGATCCAGAGCTGAGAGCTGAGGAGCTTACTTAT 1438
QY 1201 GACTTATCAAGATTTTCTTCTTCATCATCTCCAGGAGACACATGATCTGATGAT 1260
Db 1439 GACTTATCAAGATTTTCTTCTTCATCATCTCCAGGAGACACATGATCTGATGAT 1498
QY 1261 CCAAGCAGATGATGTTGCTGCTGTTGGGCTGTGTTTCTTATGTTTGGAGAC 1320
Db 1499 CCAAGCAGATGATGTTGCTGCTGTTGGGCTGTGTTTCTTATGTTTGGAGAC 1558
QY 1321 ATGAGAGAAATGCTGCTAGGTC 1344
Db 1559 ATGAGAGAAATGCTGCTAGGTC 1582

```

```

RESULT 5
LOCUS AF107834 1472 bp mRNA linear PRI 12-JUL-1999
DEFINITION Homo sapiens clone LCHI aminopeptidase mRNA, complete cds.
ACCESSION AF107834
VERSION AF107834.1 GI:5442029
KEYWORDS
SOURCE
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 1472)
AUTHORS Liu,C.H., Lin,B.Y. and Chang,L.Y.
TITLE Cloning of the human aminopeptidase gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1472)
AUTHORS Liu,C.H., Lin,B.Y. and Chang,L.Y.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia
        Sinica, RM 433, 128, Yen-Chiun-Yuan Road SEC 2, Taipei 11529, Taiwan

```

```

FEATURES
    source
        1..1472
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="LCHI"
            /tissue type="liver"
            /dev stage="fetus"
            54..1472
                /note="secretory protein"
                /codon start=1
                /product="aminopeptidase"
                /protein_id="A043214.1"
                /db_xref="GI:5442030"
                /translation="MKPLIFAFRGVHLSLCSGKALCKNGISKRTPEETKEIASCG
                    DVAKILINLAIVYKAQNRSTERYLALIVDPFLPSGSKNLEKAIQIYQWLQDDLEK
                    VHLKPIPRMVEGESAVWLEPRHKLILGSSSIGTPBEGITAEVLVTVSFDLEQ
                    RASERAGKLVVNOPIYINSRVQRTQGAANAAGVLAASLIRSVASFSYSPHG
                    IOEYOGVPRKIPACITVEDAEWMSWASHGKIVIOKMGAKATPTDSFNTVAET
                    GSKYPRQVVLVSHLSDMDVCGGAMDDGGAFTSMELSLIKLGLRPRKTLRLVMT
                    AEEQGVGAFOYIQLHKVINSYSLVMSBAGTFLPTGLDPTFSBKAALMEVMSLL
                    QPNTIQVSHSGEIDINFIQAGVPGASLLDVLRYFFPHSHGDTMTWMDPRQNV
                    AAIVVAIVSVVADMEMLRS"

```

# CDS

```

BASE COUNT 411 a 325 c 381 g 355 t
ORIGIN
Query Match 99.5%; Score 1337.6; DB 9; Length 1472;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1340; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1 AAGATGCACTCTTAAGAGACTTTTGAAGAATAAAGAAATAGCAGCTGGGA 60

```

```

Db 126 AAGATGCGATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATGCGACGCTGGA 185
Qy 61 GATGTTGCTAAAGCAATCATCACTACCTGTTTATGTTAAAGCCCAAGATCTTAT 120
Db 186 GATGTTGCTAAAGCAATCATCACTACCTGTTTATGTTAAAGCCCAAGATCTTAT 245
Qy 121 GAGGATTTGGCACTTCTGTTGATCTGTTGGAACCACTGAGTGGTCTCAAGAACTA 180
Db 246 GAGGATTTGGCACTTCTGTTGATCTGTTGGAACCACTGAGTGGTCTCAAGAACTA 305
Qy 181 GAAAAAGCCATCAAAATTTATGTACCAAACTGACAGAAATGCGCTGAGAAATTCAC 240
Db 306 GAAAAAGCCATCAAAATTTATGTACCAAACTGACAGAAATGCGCTGAGAAATTCAC 365
Qy 241 CTGAGAGCAATGAGAAATTCACCACTGAGAGAGGAGAAATGAGTGTGAGTGGAG 300
Db 366 CTGAGAGCAATGAGAAATTCACCACTGAGAGAGGAGAAATGAGTGTGAGTGGAG 425
Qy 301 CCAAGAAATTCATTAAGATAGCATCTGGGCTGTTGGACAGCATTTGGGACTCTCCAGAA 360
Db 426 CCAAGAAATTCATTAAGATAGCATCTGGGCTGTTGGACAGCATTTGGGACTCTCCAGAA 485
Qy 361 GGCATTACAGCAGAAATTTCTGTTGTTGACCTTTTCATGATGACAGAGAGGCTTCA 420
Db 486 GGCATTACAGCAGAAATTTCTGTTGTTGACCTTTTCATGATGACAGAGAGGCTTCA 545
Qy 421 GAACCAAGAGGAAATGTTGTTTATTAACCACTTAACATCACTCAAGACGCTG 480
Db 546 GAACCAAGAGGAAATGTTGTTTATTAACCACTTAACATCACTCAAGACGCTG 605
Qy 481 CAATACCGAAGCGAGGGGGCGGTGAGCTGCAAGGAGGGGGCTTTGGCATCTCTCAT 540
Db 606 CAATACCGAAGCGAGGGGGCGGTGAGCTGCAAGGAGGGGGCTTTGGCATCTCTCAT 665
Qy 541 CGATCCGTTGCTCTCTCATCTCACTGATGCTGATGCTGATGCTGATGCTGATGCT 600
Db 666 CGATCCGTTGCTCTCTCATCTCACTGATGCTGATGCTGATGCTGATGCTGATGCT 725
Qy 601 GCGCTGCCCCAAATTTCCAAAGCTGTTATTCGTTGAGAAATGCAAAATGATGCAAGA 660
Db 726 GCGCTGCCCCAAAGATTTCCAAAGCTGTTATTCGTTGAGAAATGCAAAATGATGCAAGA 785
Qy 661 ATGGCTCTCATGAGGATCAAAATTTGCTATGAGCTTAAGATGAGGAGCAAAAGCTTACCA 720
Db 786 ATGGCTCTCATGAGGATCAAAATTTGCTATGAGCTTAAGATGAGGAGCAAAAGCTTACCA 845
Qy 721 GATACGATTCCTTCAACACTGATGACAGATCACTGAGCAAAATTCAGAAAGAGTT 780
Db 846 GATACGATTCCTTCAACACTGATGACAGATCACTGAGCAAAATTCAGAAAGAGTT 905
Qy 781 GATACGATTCCTTCAACACTGATGACAGATGTTGGGCGAGGTCATGATGATGATG 840
Db 906 GATACGATTCCTTCAACACTGATGACAGATGTTGGGCGAGGTCATGATGATGATG 965
Qy 841 GGTGAGACCTTTATATATATGAGGAGCACTTCACTTATTAAGATCTTGGCTGCTGCA 900
Db 966 GGTGAGACCTTTATATATATGAGGAGCACTTCACTTATTAAGATCTTGGCTGCTGCA 1025
Qy 901 AAGAGACTCTGCGGCTGTTGCTCTGAGACTGACAGAAAGCAAGTGTGAGTGGCTTC 960
Db 1026 AAGAGACTCTGCGGCTGTTGCTCTGAGACTGACAGAAAGCAAGTGTGAGTGGCTTC 1085
Qy 961 CAGATTTATCAGTTACACAGAGTAAATTTTCCAACTACAGTCTGATGAGTCTGAC 1020
Db 1086 CAGATTTATCAGTTACACAGAGTAAATTTTCCAACTACAGTCTGATGAGTCTGAC 1145
Qy 1021 GCAAGAACTTTCTTAACCACTGAGGCTGAATTCATGAGAGTGAAGGCGAGGCGCATC 1080
Db 1146 GCAAGAACTTTCTTAACCACTGAGGCTGAATTCATGAGAGTGAAGGCGAGGCGCATC 1205
Qy 1081 ATGAGAGAGTTATGAGCTGCTGACGCCCTCAATATCACTCAGTCTGAGGCGCATGGA 1140

```

```

Db 1206 ATGAGAGAGGTTATGAGCTGTCAGAGCCCTCAATATCACTCAGTCTGAGCATGGA 1265
Qy 1141 GAAGGAGCAGACATCAACTTTTGGATCCAAAGCTGAGAGTGGCTGAGGCTTACTTAT 1200
Db 1266 GAAGGAGCAGACATCAACTTTTGGATCCAAAGCTGAGAGTGGCTGAGGCTTACTTAT 1325
Qy 1201 GACTTATACAGATTTCTTCTTCATCACTCCACAGAGAGACACATGATGATGAT 1260
Db 1326 GACTTATACAGATTTCTTCTTCATCACTCCACAGAGAGACACATGATGATGAT 1385
Qy 1261 CCAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTTCTTATGTTTGCAGAC 1320
Db 1386 CCAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTTCTTATGTTTGCAGAC 1445
Qy 1321 ATGAGAAATGCTGCTTACTGCTCC 1344
Db 1446 ATGAGAAATGCTGCTTACTGCTCC 1469

```

RESULT 6  
BD127520 1860 bp DNA linear PAT 18-SEP-2002  
LOCUS  
DEFINITION  
Primer for synthesizing full-length cDNA and use thereof.  
ACCESSION  
BD127520  
VERSION  
BD127520.1 GI:23222465  
KEYWORDS  
JP 2002017375-A/2951.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

REFERENCE  
Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.  
1 (bases 1 to 1860)  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE  
JOURNAL  
Patent: JP 2002017375-A 2951 22-JAN-2002;  
HELEX RESEARCH INSTITUTE  
COMMENT  
OS Homo sapiens (human)  
PN JP 2002017375-A/2951  
PD 22-JAN-2002  
PF 07-JUL-2000 JP 2000253172  
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO PI, YUJI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI SHINICHI KOJIMA, PI TETSUO OTSUKI, HISASHI KOGA

FEATURES  
source  
1. 1860  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 515 a 423 c 451 g 471 t  
ORIGIN  
FT CDS Location/Qualifiers  
(129)..(1544).

Query Match 99.4%; Score 1336; DB 6; Length 1860;  
Best Local Similarity 99.6%; Pct. No. 0;  
Matches 1339; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

Qy 1 AAGATGCGATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATGCGACGCTGGA 60
Db 201 AAGATGCGATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATGCGACGCTGGA 260
Qy 61 GATGTTGCTAAAGCAATCATCACTACCTGTTTATGTTAAAGCCCAAGATCTTAT 120
Db 261 GATGTTGCTAAAGCAATCATCACTACCTGTTTATGTTAAAGCCCAAGATCTTAT 320

```

QY 121 GAGCGATTGGCACTTGTGTTGATCTGTGAGACCCAGACTGAGTGGCTCCAGAACCTTA 180  
 Db 321 GAGCGATTGGCACTTGTGTTGATCTGTGAGACCCAGACTGAGTGGCTCCAGAACCTTA 380  
 QY 181 GAAAAACCCATCCAAATTAATGACCAAAACCTGACAGCAAGTGGGCTGGAGAAAGTTAC 240  
 Db 381 GAAAAACCCATCCAAATTAATGACCAAAACCTGACAGCAAGTGGGCTGGAGAAAGTTAC 440  
 QY 241 CTGGAGCCAGTGAAGATCCCACTGGAGAGGGAGAGAAATGAGTGTGATGCTGGAG 300  
 Db 441 CTGGAGCCAGTGAAGATCCCACTGGAGAGGGAGAGAAATGAGTGTGATGCTGGAG 500  
 QY 301 CCAAGAAATTCATAAGATAGCCATCTGTGGCTTGGACAGACATTTGGGACTCTCCAGAA 360  
 Db 501 CCAAGAAATTCATAAGATAGCCATCTGTGGCTTGGACAGACATTTGGGACTCTCCAGAA 560  
 QY 361 GGCATTACAGCAAGATTTGTGTGTGACCTCTTTCATGTAAGTGCAGAGAAAGGCTTCA 420  
 Db 561 GGCATTACAGCAAGATTTGTGTGTGACCTCTTTCATGTAAGTGCAGAGAAAGGCTTCA 620  
 QY 421 GAAGCAAGAGGAGATTTGTTTATTAACCACTTATCATCACTACTCAAGAGAGGCTG 480  
 Db 621 GAAGCAAGAGGAGATTTGTTTATTAACCACTTATCATCACTACTCAAGAGAGGCTG 680  
 QY 481 CAATACCGAAGCAGAGGGGCGGTGAGAGCTCCAAAGTGGGGCTTTGGCATCTCTCAT 540  
 Db 681 CAATACCGAAGCAGAGGGGCGGTGAGAGCTCCAAAGTGGGGCTTTGGCATCTCTCAT 740  
 QY 541 CGATCCGTTGCTCTCTCTCTCATCTACAGTCTCTCAACAGGTTATTCAGAAATACAGAG 600  
 Db 741 CGATCCGTTGCTCTCTCTCTCATCTACAGTCTCTCAACAGGTTATTCAGAAATACAGAG 800  
 QY 601 GGGCGGCCCAAAATTTCCAGAGCGCTGATTAACGGTGAAGATGAGAAATGATGCAAGA 660  
 Db 801 GGGCGGCCCAAAATTTCCAGAGCGCTGATTAACGGTGAAGATGAGAAATGATGCAAGA 860  
 QY 861 ATGGCTTCTCATGAGATCAAAATTTGTCATTCAGTCAAGTGAAGTGGGGCAAGACCTACCA 720  
 Db 861 ATGGCTTCTCATGAGATCAAAATTTGTCATTCAGTCAAGTGAAGTGGGGCAAGACCTACCA 920  
 QY 721 GATATGATTTCTTCAACACTGTGACAGAGATCACTGGGAGCAAAATATCCAGAACAGTT 780  
 Db 921 GATATGATTTCTTCAACACTGTGACAGAGATCACTGGGAGCAAAATATCCAGAACAGTT 980  
 QY 781 GATATGATTTCTTCAACACTGTGACAGAGATGTTGGGAGGAGCCATGAGATGATGAC 840  
 Db 981 GATATGATTTCTTCAACACTGTGACAGAGATGTTGGGAGGAGCCATGAGATGATGAC 1040  
 QY 841 GGTGAGACCTTTATATCATGAGGAAGCACTCTCACTTATTAAGATCTTGGGCTGGTCCA 900  
 Db 1041 GGTGAGACCTTTATATCATGAGGAAGCACTCTCACTTATTAAGATCTTGGGCTGGTCCA 1100  
 QY 901 AAGAGACTCTGCGGCTGTGCTCTGACTGACAGAAAGCAAGGTGAGTGTGCTTCC 960  
 Db 1101 AAGAGACTCTGCGGCTGTGCTCTGACTGACAGAAAGCAAGGTGAGTGTGCTTCC 1160  
 QY 961 CAGTATTAATCAGTTACAGAAAGTAAATTTTCCACTGCTGTGTGATGAGTCTGAC 1020  
 Db 1161 CAGTATTAATCAGTTACAGAAAGTAAATTTTCCACTGCTGTGTGATGAGTCTGAC 1220  
 QY 1021 GCAGGAACCTTTATACCACTGAGGCTGCAATTCCTGACAGTGAAGGAGGAGCCATC 1080  
 Db 1221 GCAGGAACCTTTATACCACTGAGGCTGCAATTCCTGACAGTGAAGGAGGAGCCATC 1280  
 QY 1081 ATGAGAGAGTTATGAGCTGTGACAGCCCTCAATATCACTCAGGTCTGAGCCATGGA 1140  
 Db 1281 GTGAGAGAGTTATGAGCTGTGACAGCCCTCAATATCACTCAGGTCTGAGCCATGGA 1340  
 QY 1141 GAAGGAGAGCATCAATTTTGGATCCAGCTGAGAGTGGCTGGAGCCAGTCTACTTAT 1200  
 Db 1341 GAAGGAGAGCATCAATTTTGGATCCAGCTGAGAGTGGCTGGAGCCAGTCTACTTAT 1400  
 QY 1201 GACTTAATACAGTATTTCTTCTCATCACTCCACGAGACACCATGATGTATGAT 1260

Db 1401 GACTTAATACAGTATTTCTTCTCATCACTCCACGAGACACCATGATGTATGAT 1460  
 QY 1261 CCAAGAGATGATGTTGTGCTGCTGCTGTTGGCTGTGTTTATGTTGTTGACAC 1320  
 Db 1461 CCAAGAGATGATGTTGTGCTGCTGCTGTTGGCTGTGTTTATGTTGTTGACAC 1520  
 QY 1321 ATGAGAGAAATGCTGCTAGGTCG 1344  
 Db 1521 ATGAGAGAAATGCTGCTAGGTCG 1544

RESULT 7  
 AK075132  
 LOCUS  
 DEFINITION Homo sapiens cDNA FLJ90651 f18, clone PLACB1004482, moderately similar to Rattus norvegicus hematopoietic lineage switch 2 related protein (Hl62-rp) mRNA.  
 ACCESSION AK075132  
 VERSION AK075132.1 GI:22761022  
 KEYWORDS oligo cloning; f18 (full insert sequence).  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Isegai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T., Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S., Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Kojima,S., Nagahara,K., Masuno,Y., Ono,T., Okano,K., Yoshikawa,Y., Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and Nimomiya,K.  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

TITLE NEDO human cDNA sequencing project  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1860)  
 AUTHORS Isegai,T. and Otsuki,T.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-MAR-2002) Takao Isegai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genome@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.).  
 Location/Qualifiers

#### FEATURES

source  
 1..1860  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="PLACB1004482"  
 /tissue\_type="Placenta"  
 /clone\_lib="PLACB1"  
 /note="Cloning vector: pME18SFL3"  
 129..1547  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="PAC11423.1"  
 /db\_xref="GI:22761023"  
 /translation="MKRLIFAPRGVHLLSLCSKATCRNGISRTREBEIEEIASCGDVAKATLNLAVYGAQRSTBRLLLVIVGPRUSGSKNLEKAIQINQOQDELEK VHLSEPRIPMERGESAVMLEPRIHKIALIGLSSIGTPEGIITAEVLVTSDELEQ RRSSEARGLIVVNPQPIYNSRTQVETQGAVERAKGALASLIRSVASFISYPHIG IOEXODGPKIPFACITVEDAEMMSRMAHGIKIVIOKGAATKYPTDSFNTVAEIT GSKYRQEVAVVGLSDMDVYGQGMDDGGAFTSWELSLIKDLGLPKKTLRLVMT AERGCYGAFOYVOYHRYVINSYNSIWSASAGTTLTPGLGLOFTGSEKRAIYERVMSLT QPLNITGVLSHGSTIDINFWIQGVPPASLUDDLKYFFHSHGDTMTWMDPKQNV AAAVWAVSVVAVDMEEMLPSS"

BASE COUNT 515 a 423 c 451 g 471 t  
 ORIGIN







```

/product="blood plasma glutamate carboxypeptidase
precursor"
/protein_id="AA011418.1"
/db_xref="GI:4877698"
/translation="MKFLIRFPGVHVLISCSKATCRNGISKRTPEEIKERIASG
DVAKATINLAVYKAQNRYSERLALVDVTPRLSSKNLEKAIQIMYQMLQDGLSK
VHLEPVI PHEMERGESAVMLEPRLIKHIALIGLSSIGTPREGITLAVLVTSFDEIQ
RRASEARKIVVYNOPIYNSRTVOYRTQGAVEAKYALASLIRSVAFSISPTG
IOEXODGVPKIPACTIATEDVEMSRMSAGIKIQLDKGAKTYPPTDSFNVAET
GSRYRQVAVVSGHLDSDMDYGGGAMDDGCAFTSMELSLIKDLGAPKRTLRVMT
AEEGAGVGAFOYVQLHKVNTSNVLSVESDAGTLPGLDPTSEKRAIMEEMSLI
QPLNTIYVLSHGETDINFWIQAGVPASLIDLDYKFFPHSHGDTMTWDSRWML
LILFLFLMLLQTKKCCIGPRNSKKEFTSCFWPGLIGLQMLKTPHITISNSSSK
HNSISCFLLSFLILSKFSRKRNSPLPPT"
25..96
sig_peptide
/gene="PGCP"
157..1647
mat_peptide
/gene="PGCP"
/product="blood plasma glutamate carboxypeptidase"
205..207
/misc_feature
/gene="PGCP"
/notes="putative; glycosylation site"
559..561
/misc_feature
/gene="PGCP"
/notes="putative; glycosylation site"
1081..1083
/misc_feature
/gene="PGCP"
/notes="putative; glycosylation site"
1090..1092
/misc_feature
/gene="PGCP"
/notes="putative; glycosylation site"
1210..1212
/misc_feature
/gene="PGCP"
/notes="putative; glycosylation site"
BASE COUNT 522 a 391 c 422 g 457 t 2 others
ORIGIN
Query Match 98.8%; Score 1327.2; DB 9; Length 1794;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1340; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1 AAGATGGCATCTCTAAGAGAGCTTTGAAGAAATAAAGAAATAGCCAGCTGTGA 60
Db 97 AAGATGGCATCTCTAAGAGAGCTTTGAAGAAATAAAGAAATAGCCAGCTGTGA 156
QY 61 GATGTGCTAAGCAATCATCACTAGCTGTTATGTGTAAAGCCAGAACAGATCTAT 120
Db 157 GATGTGCTAAGCAATCATCACTAGCTGTTATGTGTAAAGCCAGAACAGATCTAT 216
QY 121 GAGGATTTGGCATCTTGTGTGATATCTGTGGACCCAGACTGAGTGGCTCCAGAACCTA 180
Db 217 GAGGATTTGGCATCTTGTGTGATATCTGTGGACCCAGACTGAGTGGCTCCAGAACCTA 276
QY 181 GAAAAAGCATCCAAATATATGACCAAAACCTGACAGCAAGATGGGCTGGAGAAAGTTAC 240
Db 277 GAAAAAGCATCCAAATATATGACCAAAACCTGACAGCAAGATGGGCTGGAGAAAGTTAC 336
QY 241 CTGAGGCACTGAGAAATACCCCACTGGAGAGGGGAGAAAGATCACTGTGATCTGGAG 300
Db 337 CTGAGGCACTGAGAAATACCCCACTGGAGAGGGGAGAAAGATCACTGTGATCTGGAG 396
QY 301 CCAAGATTTCAATAGATAGCCATCTTGGCTTTGGCAGACAGATTTGGAGATCTTCCAGAA 360
Db 397 CCAAGATTTCAATAGATAGCCATCTTGGCTTTGGCAGACAGATTTGGAGATCTTCCAGAA 456
QY 361 GGCATTACAGAGAAAGTTCTGTGTGACCTTTTCAGATGAAGTGCAGAGAAAGGCTTCA 420
Db 457 GGCATTACAGAGAAAGTTCTGTGTGACCTTTTCAGATGAAGTGCAGAGAAAGGCTTCA 516
QY 421 GAAGCAAGAGGAAAGATTTGTTTATACCAACCTTACATCACTACTCAAGAGCGGTG 480
Db 517 GAAGCAAGAGGAAAGATTTGTTTATACCAACCTTACATCACTACTCAAGAGCGGTG 576

```

```

QY 481 CAATACGAAACGAGGAGGCGGTGAAAGCTGCCAAGGTGGGGCTTTGGCATCTCAT 540
Db 577 CAATACGAAACGAGGAGGCGGTGAAAGCTGCCAAGGTGGGGCTTTGGCATCTCAT 636
QY 541 CGATCCGTGGCTCTTCTTCCATCTACAGTCTCTCAACAGATTTACAGAAATCCAGAT 600
Db 637 CGATCCGTGGCTCTTCTTCCATCTACAGTCTCTCAACAGATTTACAGAAATCCAGAT 696
QY 601 GGGGTGCCCAAAATTCACACAGCTGATTTAGCGTGAAGATCAGAAATGATGTCAAG 660
Db 697 GGGGTGCCCAAAATTCACACAGCTGATTTAGCGTGAAGATCAGAAATGATGTCAAG 756
QY 661 ATGGCTTTCATGAGATCAAAATTTGTTCATTCAGCTAAAGATGGGGCAAAAGCTTCCA 720
Db 757 ATGGCTTTCATGAGATCAAAATTTGTTCATTCAGCTAAAGATGGGGCAAAAGCTTCCA 816
QY 721 GATATGATTTCTTCAACACTGTAGACAGATCACTGGAGCAAAATATCCAGAACAGTT 780
Db 817 GATATGATTTCTTCAACACTGTAGACAGATCACTGGAGCAAAATATCCAGAACAGTT 876
QY 781 GTACTGGTCACTGAGACATCTGACACAGTGGATTTGGGCGAGGTGCANTGATATGGC 840
Db 877 GTACTGGTCACTGAGACATCTGACACAGTGGATTTGGGCGAGGTGCANTGATATGGC 936
QY 841 GGTGAGGCTTTATATCATAGGAGACATCTCACTTATTAAGATCTTGGGCTGCTCCA 900
Db 937 GGTGAGGCTTTATATCATAGGAGACATCTCACTTATTAAGATCTTGGGCTGCTCCA 996
QY 901 AAGAGACTCTGCGGCTGTGTCTGTGACTGCAAGAAAGACAGGTGAGTTGTGCTTC 960
Db 997 AAGAGACTCTGCGGCTGTGTCTGTGACTGCAAGAAAGACAGGTGAGTTGTGCTTC 1056
QY 961 CAGTATTAACATTAACAACAAGATTAATTTCCATCACTGCTGTGATGAGTCTGAC 1020
Db 1057 CAGTATTAACATTAACAACAAGATTAATTTCCATCACTGCTGTGATGAGTCTGAC 1116
QY 1021 GCAGGAACCTTCTTACCACTGAGGCTGCAATTCATCTGACAGTAAAGAGGCGAGCCATC 1080
Db 1117 GCAGGAACCTTCTTACCACTGAGGCTGCAATTCATCTGACAGTAAAGAGGCGAGCCATC 1176
QY 1081 ATGAGAGAGTTATGAGCTGTGACAGCCCTCAATATCACTCAGGTCTGTAGCCATGGA 1140
Db 1177 ATGAGAGAGTTATGAGCTGTGACAGCCCTCAATATCACTCAGGTCTGTAGCCATGGA 1236
QY 1141 GAAGGACACACATCAACTTTTGGATTCAGAGCTGAGTGGCTTGGAGCACTTACTTGAT 1200
Db 1237 GAAGGACACACATCAACTTTTGGATTCAGAGCTGAGTGGCTTGGAGCACTTACTTGAT 1296
QY 1201 GACTTATACAAATATTTCTTCCATCACTCCAGAGAGACCAATGATCTGATGAT 1260
Db 1297 GACTTATACAAATATTTCTTCCATCACTCCAGAGAGACCAATGATCTGATGAT 1356
QY 1261 CCAAGCAGATGATGTTGTGCTGTGTTGGGCTGTGTTCTTATGTTGTGACAGC 1320
Db 1357 CC-CAGCAGATGATGTTGTGCTGTGTTGGGCTGTGTTCTTATGTTGTGACAGC 1415
QY 1321 ATGGAAGAAATGCTGCTAGGTC 1344
Db 1416 ATGGAAGAAATGCTGCTAGGTC 1439

```

```

RESULT 9
BC037067 1796 bp mRNA linear ROD 16-APR-2003
LOCUS
DEFINITION
MUS musculus plasma glutamate carboxypeptidase, mRNA (cDNA clone
MGC:46857 IMAGE:4989316), complete cds.
ACCESSION
BC037067
VERSION
BC037067.1 GI:22477497
KEYWORDS
MGC.
SOURCE
MUS musculus (house mouse)
ORGANISM
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```



```

QY 781 GTACTGTCAGTGCACATCTGACAGCTGGAGATGTTGGGACGGGTGCATGATGATGAC 840
DB 909 GTCTGTGTCAGTGCACATCTGACAGCTGGAGATGTTGGGACGGGTGCATGATGATGAC 968
QY 841 GGTGGAGCCTTTATATATATGAGGAAAGCACTCTCACTTATTAAGATCTTTGGGCTGCTCA 900
DB 969 GGTGGAGCCTTTATATATATGAGGAAAGCACTCTCACTTATTAAGATCTTTGGGCTGCTCA 1028
QY 901 AAGAGAGCTGCGGGCTGGTGGTCTGAGCTGCGAAGAAAGCAAGGTGAGTGTGGGCTTC 960
DB 1029 AAGAGAGCTGCGGGCTGGTGGTCTGAGCTGCGAAGAAAGCAAGGTGAGTGTGGGCTTC 1088
QY 961 CAGTATATCAGTATCAAGAGTAAATATTTCCAACTACAGTCTGTGATGAGTCTGAC 1020
DB 1089 CAGTATATCAGTATCAAGAGTAAATATTTCCAACTACAGTCTGTGATGAGTCTGAC 1148
QY 1021 GCAGGAACCTTTCTTACCCACTGGGCTGCAATTCAGTGGCAGTGAAGGCGGCGCATC 1080
DB 1149 TCAGGAACCTTTCTTACCCACTGGGCTGCAATTCAGTGGCAGTGAAGGCGGCGCATC 1208
QY 1081 ATGAGAGAGGTTATGAGCCTGGCTGCGAAGCCCTCAATATCACTGAGTCTGAGCCATGGA 1140
DB 1209 ATGAGAGAGGTTATGAGCCTGGCTGCGAAGCCCTCAATATCACTGAGTCTGAGCCATGGA 1268
QY 1141 GAAAGGACAGACATCACTTTTGGATCCAAAGCTGAGTGCCTGGAGCCAGTCTACTGAT 1200
DB 1269 GAAAGGACAGACATCACTTTTGGATCCAAAGCTGAGTGCCTGGAGCCAGTCTACTGAT 1328
QY 1201 GACTTATACAGATATTTCTTTCATCACTCCACGAGAGACACATGACTGTGATGAT 1260
DB 1329 GACTTATACAGATATTTCTTTCATCACTCCACGAGAGACACATGACTGTGATGAT 1388
QY 1261 CCAAGAGAGATGATGTTGCTGCTGCTGTTGGGCTGTGTTTATATGTTGTCAGAC 1320
DB 1389 CCAAGAGAGATGATGTTGCTGCTGCTGTTGGGCTGTGTTTATATGTTGTCAGAC 1448
QY 1321 ATGAGAGAAATGCTGCTAGGTCC 1344
DB 1449 ATGAGAGAAATGCTGCTAGGTCC 1472

```

RESULT 10  
AF107835 1596 bp mRNA linear ROD 12-JUL-1999  
LOCUS AF107835  
DEFINITION Mus musculus c1one LCH-M1 aminopeptidase mRNA, complete cds.  
ACCESSION AF107835  
VERSION AF107835.1 GI:5442031  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS Liu, C.H., Lin, B.Y. and Chang, L.Y.  
TITLE Cloning of the mouse aminopeptidase gene  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1596)  
AUTHORS Liu, C.H., Lin, B.Y. and Chang, L.Y.  
TITLE Direct Submission  
JOURNAL Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia  
Sinica, RM 433, 128, Yen-Chiun-Yuan Road SEC 2, Taipei 11529, Taiwan

FEATURES  
source  
1..1596  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="LCH-M1"  
/tissue\_type="liver"  
184..1596  
/note="secretory protein"  
/codon\_start=1  
/product="aminopeptidase"  
/protein\_id="A0A3215.1"

CDS

184..1596  
/note="secretory protein"  
/codon\_start=1  
/product="aminopeptidase"  
/protein\_id="A0A3215.1"

```

/db_xref="GI:5442032"
/translation="MRSLEFLFIVHLALGSGKAVPKNGVSRPREIKELIANVEDY
ARAIALAVAYKQNSYERLGLINDVBPRLSGSKNLEKALIDMIONLODLELVH
LEQVRIPIHMERGESAVMLEPRIRHKAIIIGLSSIGTPEGITAEVLVASFELQR
ASEBARKIIVYNOPPYTGEKTVQVRQVGALEAKVAASLIQSVASFISHTGIQ
KODGVPKIPIACIIVEDAEEMSRMSRGRKIVIHDEMKAKTYPDPSFNTVEITGS
MYEEVAVLSGLHSDVQALDDGGAFISHEALSLVLDLRPRTRLRLVLMFAE
BOGIGASOYVELHKANISKYSLVMEADSGTFLPTGIQFTGSKARAIKMEVNLLOP
LNTYKPSNEGCDINPMIOAGVPGASLBDLKYKFFPHSHSDPTVMDPKMNVA
AVAAVAIVVADMDENLPRS"
BASE COUNT 415 a 357 c 421 g 403 t
ORIGIN
Query Match 80.7%; Score 1084.8; DB 10; Length 1596;
Best Local Similarity 87.9%; Pred. No. 2.4e-297;
Matches 1182; Conservative 0; Mismatches 162; Indels 0; Gaps 0;
QY 1 AAGATGGCATCTCTAAGAGACTTTTGAAGAAATAAAGAAATAGCAGCTGTGA 60
DB 250 AAGATGGTGTCTCTAGCGACATTTGAGAAATAAAGAAATAGCAGTATGAA 309
QY 61 GATGTTGCTAAAGCATCATCAACTAGCTGTTTATGTTAAAGCCGAAACAGTCTAT 120
DB 310 GATGTTGCTAAAGCAATTATCAACTGTCTGTTATGTTAAATACCGAACCGTCTAT 369
QY 121 GAGCGATTGGCATTCTGTTGTTGATCTGTTGACCCGACGATGAGTCTCCAGAACCTA 180
DB 370 GAGCGTTGGACCTCTAGTTGATACGTGTGACCCGACGATGAGTCTCTAAGAACCTA 429
QY 181 GAAAGAGCCATCCAAATTTATGTACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTTAC 240
DB 430 GAGAAAGCTATTCAAATCATGTACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTTAC 489
QY 241 CTGAGCCAGTGAAGATATCCCACTGGGAGAGGGAGAGAAATCAGTGTGATGCTGAG 300
DB 490 CTGAGCAGGTCAAGAAATACCCCACTGGGAGAGGGAGAAATCAGTGTGATGCTGAG 549
QY 301 CCAAGAAATTCATTAAGATAGGCATCCGCGGCTTGGCAGCAGCAATGGGACTCCGCCGAA 360
DB 550 CTTGAAATTCACAGAGTGGCTATTTAGTTGGCAGCAGCAATGGGACTCCGCCGAA 609
QY 361 GGCATTACAGCAGAAAGTTCTGTGTGTGACCTCTTTGATGAACTGCAGAGAGGCGCTCA 420
DB 610 GGCATTACAGCAGAAAGTGTGTGTGTGACCTCTTTGATGAACTGCAGAGAGGCGCTCA 669
QY 421 GAAAGAGAGGAGATGTTGTTTATTAACCAACTTATCAACTAATCTCAAGACGCTG 480
DB 670 GAAAGAGAGGAGATGTTGTTTATTAACCAACTTATCAACTAATCTCAAGACGCTG 729
QY 481 CAATACCGAAGCAGAGGGGGGGGAGAGCTGCCAAGGTGGGGGCTTGGCATCTCATAT 540
DB 730 CAGTACCGGGGTGAGGGAGGTGTGAAAGCTGCCAAGGTGGAGGTGTGGCATCTCATATC 789
QY 541 CGATCCGTGGCTCTCTCTCATCTACAGTCTCAACAGATATTCAGAAATACAGAGAT 600
DB 790 CAATCAGTACGTTCTTTTCAATCTACAGTCTCAACAGATATTCAGAAATACAGAGAT 849
QY 601 GGCCTGCCAAATTCACAGCCTGTATTACGGTGAAGATGCAAAATGATGTCAAGA 660
DB 850 GGTGTGCCCAAGATTCACAGCCTGTATACAGTGAAGATGCAAAATGATGTCTCA 909
QY 661 ATGCTTCTCATGGAGTCAAAATTTGTATCTCACTAAGTAAAGTGGGGCAAGACTACCA 720
DB 910 ATGCTTCTGTGTGAACAAATTTGTATCTTGTGAAGTGGAGCAAGACTATCA 969
QY 721 GATACGATTCCTTCAACACTGTAGCAGATCACTGGAGCAAAATATTCAGAAAGGTT 780
DB 970 GATACGATTCCTTCAACACTGTGTGAGAGATCACTGGAGCAATGATATTCAGAGAGGTT 1029
QY 781 GTACTGTCAGTGCACATCTGACAGCTGGAGATGTTGGGACGGGTGCATGATGATGAC 840
DB 1030 GTCTGTGTCAGTGCACATCTGACAGCTGGAGATGTTGGGACGGGTGCATGATGATGATG 1089

```

QY	841	GGTGGAGCCTTATATATCATNCGGGAACACCTCCACTATATTAAGAATCTGGGCTGGCTCA	900
Db	1090	GGTGGAGCCTTATATATCATNCGGGAACACCTCTACTGTGTAAGAATCTGGGCTGGCTCA	1149
QY	901	AAGAGACTCTGGCGGCTGGTGTCTGGACTGCGAAGAACAAGGTGAGTGTGTCTTC	960
Db	1150	AAGAGACTCTGGCGGCTGGTGTCTGGACTGCGAAGAACAAGGAGGAATGTGTCTTC	1208
QY	961	CAGTATTTACAGTTACACAGGTAATAATTCCATCTACGTCTGGTATGGAAGTCAAC	1020
Db	1210	CAGTATTTACAGTTACATBAGGCAATATTTCCAACTACAGTTGTGTATGTGAAGCTAAC	1266
QY	1021	GCAGAACTTCTTACCCACTGGGGCTGCATTCACGTGGCAGTGAAGAAGCCAGGSCATC	1080
Db	1270	TCAGAACTTCTTACCCACTGGACTGCAGTTCAACGGGAGTGAACAAGCCAGGSCATTC	1322
QY	1081	ATGGAGAGGTTATBAGAGCTGTGTGAGAGCCCTCATATACCTCAGSTCTCTAGCCATGGA	1144
Db	1330	ATGAAGAAGTATATTAATCTTGTGAAACCCCTCAATGTGCACCAAGSTCTTTAGTAAATGA	1388
QY	1141	GAAGGACAGACATCAACTTTTGGATCCAAAGCTGAGAGTCCCTGAGACCAAGTCTACTTAT	1200
Db	1390	GAAGGAATGAAATTTAACTTCTTGATCCAAAGCTGAGATCCTGAGGCCAGTCTGTGAT	1449
QY	1201	GACTTATATCAAGTATTTCTTCTTCATCACTCACTCCACGAGACACATGACTGTCAATGAT	1266
Db	1450	GACTTGTATCAAGTATTTCTTCTTCATCACTCCACGAGACACACATGACTGTCAATGAT	1508
QY	1261	CCAAAGCAGATGAATGTTGCTCTCTGTGTTGGGCTGTGTTCTTATATGTTGTTCAAC	1320
Db	1510	CCAAAGCAGATGAATGTTGCTCTCTGTGATGAGGCTGTGTTGCTTATATGTTGTTCAAC	1566
QY	1321	ATGGAAGAAATGCTGCTTAGGTCC	1344
Db	1570	ATGATGAATGCTGCTTAGGTCC	1593

```

RESULT 11
AF131077
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS

```

Query Match	Local Similarity	Score	DB	Length
Base Count	474 a	385 c	428 g	439 t
ORIGIN				
Query Match	78.9%	Score 1060.8	DB 10	Length 1726
Local Similarity	86.8%	Pred. No. 1.6e-250		
Matches 1167	Conservative 0	Mismatches 177	Indels 0	Gaps 0
<pre> /note="involved in liver regeneration; contains unique amexin domain at the N-terminus" /codon_start=1 /product="liver amexin-like protein" /protein_id="AAF6518.1" /db_xref="gi:7108713" /translation="MRPLFLPVAVVHLPSLSGSKAIYKSVGORTFPOSTKEIANE DVAKAIINLAVYKIQNRSYERIGLVDITPPLSSKULEKAIQIQMYLQODGLEN VHLAEQVNI PMHERGEEBSAVNVPR IHLIALILGIGSTPEBISITAEVLVASFVEIG RRLSEARGKILVNVNOPYDTQVOYHERBAVAKGASVLSIRSVASISPHIG HOGQEDVPKIPACITIEDAEMMSRTASGDKIVHLKMGATPEPTDSIFNVALET GSKRPEEVPLVSGHLSMDVQGLDDGGAPFISWELATSLXVQGLRKPRTLPLVMT AERGGGASQSYEIEHKANISKYSIYMEALSISGFLPLPGIOFTSDSRALMKKWSLL OPLATIKVENDAGTDINFIWIOGVPRASLRDLYKTFPHSHSGDTMTAMDYKQNV AAAVVAVAVAVAVDMSEMLPRS" </pre>				
1	AAGATGCGCATCTCTTAAGAGCACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGA	60		
135	AAGAGTGCTGTTTCTGCGCAACCTTTCAGAAATATAAGAAATAGCCACTATGAA	194		
61	GATGTTCTAAGCAATCATCAACCTAGCTGTTTATGTAAGCCAGAACATCTTAT	120		
195	GATGTTCTAAGCAATCATCAACCTGCTGTTTATGTAAGAAATACAGAACGGTCTAT	254		
121	GAGCGATTGGCACTTCTGGTTGATACCTGTTGACCCAGACTGAGTGGCTCCAGAACCTA	180		
255	GAGGTTTGGACCTTCACTGATGTAACCTGTTGACCCAGACTGAGTGGCTCTTAAGAACCTA	314		
181	GAAAGCGCATCCAAATTTATGTATACCAAACTGCGAGAAAGTGGCTGAGAAAGTTCCAC	240		
315	GAGAAAGCTATCCAAATCATGTACCAAACTGCGAGAAAGTGGCTGAGAAAGTTCCAC	374		
241	CTGAGCGCATGAGAAATATCCCACTGCGAGAGGGGAGAGAAATACAGCTGTATGCTGGAG	300		
375	CTGAGCGCATGAGAAATATCCCACTGCGAGAGGGGAGAGAAATATGCAAGTATGCTGTG	434		
301	CCAAGATTCATTAAGATAGCCATCTTGGGTTTGGCAGCAGCATTTGGGACTCTCTCCAGAA	360		
435	CCTCGAATTCACAAAGTTGGCTATTTTGAAGGCTTGGCGGAGCATTTGGGACTCTCTCCAGAA	494		
361	GGCATTTACAGAGAAAGTTCTGCTGGTGAACCTCTTTGATGAACCTGAGAGAGAGGCTCTCA	420		
495	GGTATCACAGAGAAAGTCTGCTGGTGAACCTCTTTGATGAACCTGAGAGAGAGGCTCTCA	554		
421	GAAAGCGAGGAAAGATTTGTTTATTAACCACTTATCACTACTCAAGAGCGGTG	480		
555	GAGCGAAGAGGAAAGATTTGTTTATTAACCACTTATCACTACTCAAGAGCGGTG	614		
481	CAATACCGAAGCGCAGGGGGCGGTGGAAGCTGCCAAGTGGGGCTTTGGCATCTCTCAAT	540		
615	CAGTACCGGGAGCGCGGAGCTGTGGAAGCTGCCAAGTGGGGGCGCTGTGGCATCTCTCATC	674		
541	CGATTCGGGCGCTCTCTCATCTTAAGTCTTACACAGGTTATTCAGAAATACAGAGAT	600		
675	CGATTCAGAGTCTCTTTTCCATCTTAACAGTCTTACACAGGTTATTCAGAAATACAGAGAT	734		
601	GGCGTGGCCAAATTTCCACAGCCTGTATTAACGTTGGAAGATGAGAAATGATCTCAAGA	660		
735	GGTGTGCCCAAGATTTCCACAGCCTGTATTAACCAATTAAGATGAGAAATGATCTCTCGA	794		
661	ATGCGCTTCTCAATGAGATCAAAATTTGATTCAGCTAAGATGGGGGCAAAAGCTTACCCA	720		
795	ACGGCTTCTGTTGGGAGCAAAATTTGATTCATCTGAAATGGGAGCAAAAGCTTATCCA	854		
721	GATACGTATTTCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGATT	780		
855	GATACGTATTTCTTCAACACTGTGTGACAGAGATCACTGGAGCAAAATATCCAGAACAGATT	914		

QY 781 GTATGTCAGTGCATCTGAGCAGCTGGATGTTGGGAGGGTGCATGATGATGCG 840  
 Db 915 GTCTGTCAGTGCATCTGAGCAGCTGGAGCTGGGAGGGTGCATGATGATGCG 974  
 QY 841 GGTGAGACCTTTATATATATGAGGAGCACTCTCATTTAATAAGTCTTGGCTCGTCA 900  
 Db 975 GGTGAGACCTTTATATATATGAGGAGCACTCTCATTTAATAAGTCTTGGCTCGTCA 1034  
 QY 901 AAGAGACTCTGGGGCTGTGCTCTGAGCTGAGAGCAAGAGTGAAGTGGCTTC 960  
 Db 1035 AAGAGACTCTGGGGCTGTGCTCTGAGCTGAGAGCAAGAGTGAAGTGGCTTC 1094  
 QY 961 CAGTATTATGAGTACCAAGATTAATTTTCAACTACTGTGTGATGAGTCTGAC 1020  
 Db 1095 CAGTATTATGAGTACCAAGATTAATTTTCAAGTACAGTTTGTGATGAGTCTGAC 1154  
 QY 1021 GCAGGAACCTTTTACCACTGGCTGCAATTCATCTGCAAGTGAAGGCGGCAATC 1080  
 Db 1155 TCAGGAACCTTTTACCACTGGCTGCAATTCATCTGCAAGTGAAGGCGGCTATC 1214  
 QY 1081 ATGAGAGAGTTATGAGCTGCTGAGCCCTCAATATCACTGAGTCCGAGGCATGGA 1140  
 Db 1215 ATGAGAGAGTTATGAGCTGCTGAGCCCTCAATATCACTGAGTCCGAGGCATGGA 1274  
 QY 1141 GAAAGGACAGACATCACTTTTGGATCCAGCTGAGTGGAGCCAGTCTACTTGAAT 1200  
 Db 1275 GAAAGGACAGACATCACTTTTGGATCCAGCTGAGTGGAGCCAGTCTACTTGAAT 1334  
 QY 1201 GACTTATCAAGATTTCTTCTTCATCACTCCACGAGAGACACATGATCTGATGAT 1260  
 Db 1335 GACTTATCAAGATTTCTTCTTCATCACTCCACGAGAGACACATGATCTGATGAT 1394  
 QY 1261 CCAAGAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
 Db 1395 CCAAGAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1454  
 QY 1321 ATGAGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1344  
 Db 1455 ATGAGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1478

RESULT 12  
 AX400829 1778 bp DNA linear PAT 06-JUN-2002  
 LOCUS AX400829 505 from Patent WO0210453.  
 DEFINITION AX400829  
 ACCESSION AX400829.1 GI:21337009  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE  
 AUTHORS Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and  
 1. Blaschke, M.R.  
 TITLE Molecular toxicology modeling  
 JOURNAL Patent: WO 0210453-A 505 07-FEB-2002;  
 Gene Logic, Inc. (US)  
 FEATURES  
 source  
 1. 1778  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /note="EMBL/GenBank Accession No. AF097723"  
 BASE COUNT 488 a 391 c 442 g 457 t  
 ORIGIN

Query Match 78.9%; Score 1060.8; DB 6; Length 1778;  
 Best Local Similarity 86.8%; Pred. No. 1.6e-290;  
 Matches 1167; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 1 AAGATGGCATCTTAAGAGACTTTTGAAGAAATTAAGAGAAATAGCCAGTGTGA 60  
 Db 193 AAGATGGTGTCTTCTGAGCAACATTTCAAGAAATTAAGAGAAATAGCCATATGAA 252  
 QY 61 GATGTCCTTAAGCAATCATCAACTAGCTGTTTATGTGTAAGCCCAAGATCTTAT 120  
 Db 253 GATGTCCTTAAGCAATCATCAACTAGCTGTTTATGTGTAAGAAATACCAAGAACCGGTCTAT 312  
 QY 121 GAGGATTTGGCACTTCGTTGATGATCTGTGAGCCGAGCTGAGTGGCTCAAGAACTTA 180  
 Db 313 GAGGATTTGGCACTTCGTTGATGATCTGTGAGCCGAGCTGAGTGGCTCTTAAGAACTTA 372  
 QY 181 GAAAGACCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240  
 Db 373 GAGAAACCTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 432  
 QY 241 CTGAGAGCAAGTGAATATCCCACTGGGAGAGGGGAGAGATATCAGTGTGATGCTGAG 300  
 Db 433 CTGAGAGCAAGTGAATATCCCACTGGGAGAGGGGAGAGATATCAGTGTGATGCTGAG 492  
 QY 301 CCAAGATTCATTAAGATAGGCACTCTGGGCTTGGGAGAGAGATGAGGATCTCTCCAGAA 360  
 Db 493 CTGGAATTCATTAAGATAGGCACTCTGGGCTTGGGAGAGAGATGAGGATCTCTCCAGAA 552  
 QY 361 GGCATTACAGCAGAGATTTCTGTGTGACCTCTTTCATGATGATGAGAGAGGCTCTCA 420  
 Db 553 GGTATACACAGAGATTTCTGTGTGACCTCTTTCATGATGATGAGAGAGGCTCTCA 612  
 QY 421 GAGCAAGAGAGAGATTTGTTTATTAATCAATCTTAATCAATCTTAATCAAGAGAGGCTG 480  
 Db 613 GAGCAAGAGAGAGATTTGTTTATTAATCAATCTTAATCAATCTTAATCAAGAGAGGCTG 672  
 QY 481 CAATACCGAAGCAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540  
 Db 673 CAGTACCGAAGCAGAGAGGAG 732  
 QY 541 CGATCCGTCGCTCTCTCTCATCTAGTCTCAACAGATTTTCAAGATTAACAGAT 600  
 Db 733 CGATCCGTCGCTCTCTCTCATCTAGTCTCAACAGATTTTCAAGATTAACAGAT 792  
 QY 601 GGGGTGCCCAAAATTCACACAGCTGTATTAAGGTGAGAGATGAGATGATCTCAAG 660  
 Db 793 GGGGTGCCCAAAATTCACACAGCTGTATTAAGGTGAGAGATGAGATGATGATCTCAAG 852  
 QY 661 ATGCTTCTCATGAGATCAAAATTTGATCTAGCTTAAGATGAGGAGAAAGCTTACCA 720  
 Db 853 ATGCTTCTCATGAGATCAAAATTTGATCTAGCTTAAGATGAGGAGAAAGCTTACCA 912  
 QY 721 GATCTGATTCCTTCAACACCTGTAGCAGAGATCACTGGAGCAAAATTAACAGAT 780  
 Db 913 GATCTGATTCCTTCAACACCTGTAGCAGAGATCACTGGAGCAAAATTAACAGAT 972  
 QY 781 GTATGTCAGTGCATCTGAGCAGCTGGAGATGTTGGGAGGGTGCATGATGATGCG 840  
 Db 973 GTCTGTCAGTGCATCTGAGCAGCTGGAGATGTTGGGAGGGTGCATGATGATGCG 1032  
 QY 841 GGTGAGACCTTTATATATATGAGGAGCACTCTCATTTAATAAGTCTTGGCTCGTCA 900  
 Db 1033 GGTGAGACCTTTATATATATGAGGAGCACTCTCATTTAATAAGTCTTGGCTCGTCA 1092  
 QY 901 AAGAGACTCTGGGGCTGTGCTCTGAGCTGAGAGAAAGAGAGTGAAGTGGCTTC 960  
 Db 1093 AAGAGACTCTGGGGCTGTGCTCTGAGCTGAGAGAAAGAGAGTGAAGTGGCTTC 1152  
 QY 961 CAGTATTATGAGTACCAAGATTAATTTTCAACTACTGTGTGATGAGTCTGAC 1020  
 Db 1153 CAGTATTATGAGTACCAAGATTAATTTTCAAGTACAGTTTGTGATGAGTCTGAC 1212  
 QY 1021 GCAGGAACCTTTTACCACTGGCTGCAATTCATCTGCAAGTGAAGGCGGCAATC 1080  
 Db 1213 TCAGGAACCTTTTACCACTGGCTGCAATTCATCTGCAAGTGAAGGCGGCAATC 1272  
 QY 1081 ATGAGAGAGTTATGAGCTGCTGAGCCCTCAATATCACTAGTCTGAGGCATGGA 1140

Db 1273 ATGAAGAAAGTCATGAGTCTCTGCAACCCCTCATATCACCAAGGTCTTTAATGATGCA 1332  
Qy 1141 GAAGGACAGACATCACTTTTGATCCAGGTGAGTGGCTGGAGCCAGTCTACTTAT 1200  
Db 1333 GAAGAACTGACATTAACCTTCTGATCCAGGTGAGTGGCTGGAGCCAGTCTGAGAT 1392  
Qy 1201 GACTTATCAAGTATTTCTTCTTCATCACTCCAGGAGACACCATGATGATGAT 1260  
Db 1393 GACTTATCAAGTATTTCTTCTTCATCACTCCAGGAGACACCATGATGATGAT 1452  
Qy 1361 CCAAGCAGATGATGTTGCTGCTGCTGTTGGCTGTTGTTCTTATGTTGTCAGAC 1320  
Db 1453 CCAAGCAGATGATGTTGCTGCTGCTGTTGGCTGTTGTTGCTTATGTTGTCAGAC 1512  
Qy 1321 ATGAGAAATGCTGCTGATGCTC 1344  
Db 1513 ATGAGAAATGCTGCTGATGCTC 1536

RESULT 13  
AF097723 1778 bp mRNA linear ROD 10-NOV-1998  
LOCUS Rattus norvegicus hematopoietic lineage switch 2 related protein  
DEFINITION (H1s2-rp) mRNA, complete cds.  
ACCESSION AF097723  
VERSION AF097723.1 GI:3851631  
KEYWORDS  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;  
Rattus;  
1 (bases 1 to 1778)  
Chen, Y. and Talmage, D.  
DIRECT SUBMISSION  
Submitted (07-OCT-1998) Institute of Human Nutrition, Columbia  
University, 701 West 168th Street Room 5-503, New York, NY 10032,  
USA

FEATURES  
source Location/Qualifiers  
1..1778  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Fisher"  
/db\_xref="taxon:10116"  
/cell\_line="F111"  
/cell\_type="fibroblast"  
/dev\_stage="embryo"  
1..1778  
/gene="H1s2-rp"  
121..1539  
/gene="H1s2-rp"  
/codon\_start=1  
/product="hematopoietic lineage switch 2 related protein"  
/protein\_id="AAC72384.1"  
/db\_xref="GI:3851632"  
/translation="KRFLPLFVAVVHLPSLGSKATYKSGVSGTPEIKELANVE  
DVAKIINLAVYKQNSYERLGLVTVGRLSGSKNLEALIDIMQNIQDLEN  
VHLDRVRIPIHWGRBESAVMVPRILHKAIIIGLGSISITPEGIATVAVLVASFEVLQ  
RRASERAKIIVVNPYDYGTVQYRRGRGAFAKAVASVLSVSFISYSPHTG  
HOGYODGVPIPTACTITIEDAMMSRMSRGDKIVILHMGAKTVPDTSFNTVAET  
GSKYDEEVVLVSGHDSMDVGAGALDDGGAFLISWEALSLVMDGLRKRRLRIVLMT  
ABEGGVGASQYELHKNKISYSLVMEADSGTFLPTGLQPTGSKAKAIKEVMSLL  
QPLNTKVFNDSEGTIDINFWIQGVPGASLRDLTKYFFPHSHSDTMTAMDPKOMNV  
AAAWAVVAVVADMEMLPRS"

BASE COUNT 488 a 391 c 442 g 457 t  
ORIGIN

Query Match 78.9%; Score 1060.8; DB 10; Length 1778;  
Best Local Similarity 86.8%; Pred.No.1.6e-290;  
Matches 1167; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

Qy 1 AAGAAATGCAATCTTAAGAGACCTTTTGAAGAAATAAAGAAATAGCCAGCTGTGGA 60

Db 193 AAGAGTGTGTTTCTCAGCCAACTTTCAAGAAATAAAGAAATAGCCATATGAA 252  
Qy 61 GATTTGCTAAAGCAATCATCAACCTAGCTGTTTATGTAAGCCGAGAAAGATCTAT 120  
Db 253 GATTTGCTAAAGCAATCATCAACCTAGCTGTTTATGTAAGAAATAGCCAGATCTAT 312  
Qy 121 GAGGATTTGCACTTCTGTTGATGATGTTGGAAGCCAGACTGATGAGTCCAGAACTTA 180  
Db 313 GAGGATTTGGAATCTTCTGATGATGATGTTGATGATGTTGGAAGCCAGACTGATGAGTCC 372  
Qy 181 GAAAGACATCCAAATATGATCAAAACCTGACAGAGATGAGTCCAGAAAGTTCA 240  
Db 373 GAGAAAGTATCCAAATCATGATCAAAACCTGACAGAGATGAGTCCAGAAAGTTCA 432  
Qy 241 CTGAGCCAGTGAAGATATCCCACTGAGAGAGAGAGAGAGATGATGATGATGATG 300  
Db 433 CTGAGAGAGTCAAGATATCTCACTGAGAGAGAGAGAGAGATGATGATGATGATG 492  
Qy 301 CCAAGAAATTCATAGATAGCATCTGAGTCTTGGAGAGCATTTGGGACTCTCCAGAA 360  
Db 493 CTGGAATTCAGATTTGCTATTTAGGCTTGGCGGACATTTGGGACTCTCTGAA 552  
Qy 361 GGCATTCACAGAGAGTCTGAGTCACTCTTTCATGATGATGATGATGATGATGATG 420  
Db 553 GGTATCACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 612  
Qy 421 GAAGCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
Db 613 GAGCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 672  
Qy 481 CAATACCGAAG 540  
Db 673 CAGTACCGAG 732  
Qy 541 CGATCCGAG 600  
Db 733 CGATCAG 792  
Qy 601 GCGTCCGAG 660  
Db 793 GGTGCGCAG 852  
Qy 661 ATGCTTCTCATGAG 720  
Db 853 ATGCTTCTCATGAG 912  
Qy 721 GATACGATCTCTTCAACACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
Db 913 GATACGATCTCTTCAACACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 972  
Qy 781 GATACGATCTCTTCAACACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840  
Db 973 GTCCTGTCAGTGAACATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1032  
Qy 841 GGTGAGGCTTATATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
Db 1033 GGTGAGGCTTATATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1092  
Qy 901 AAGAGAGCTGCGAGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
Db 1093 AAGAGAGCTGCGAGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1152  
Qy 961 CAGTATTCATGATTCACAGAGATATATTTTCCACTACAGTCTGATGATGATGATG 1020  
Db 1153 CAGTATTCATGATTCACAGAGATATATTTTCCACTACAGTCTGATGATGATGATG 1212  
Qy 1021 GCAGAACTCTTATACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
Db 1213 TCAGAGAACTTCTTACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1272  
Qy 1081 ATGAGAGAGTATGAG 1140

Db 1273 ATGAGAGAGTATGATCTCTCCGAAACCCCTCAATATACCAAGTCTTTATGATGCA 1332  
Qy 1141 GAAGGACAGACATCACTTTTGGATCCAGCTGAGTGCCTGGAGCAGTCTACTGAT 1200  
Db 1333 GAAGGACAGACATCACTTTTGGATCCAGCTGAGTGCCTGGAGCAGTCTACTGAT 1332  
Qy 1201 GACTTATACAGATATTTCTTTCATCACTCCACGAGACACACATGATGTCATGAT 1260  
Db 1393 GACTTATACAGATATTTCTTTCATCACTCCACGAGACACACATGATGTCATGAT 1452  
Qy 1261 CCAAGAGATGATATTTCTTTCATCACTCCACGAGACACACATGATGTCATGAT 1330  
Db 1453 CCAAGAGATGATATTTCTTTCATCACTCCACGAGACACACATGATGTCATGAT 1512  
Qy 1321 ATGAGAGAGTATGATCTCTCCGAAACCCCTCAATATACCAAGTCTTTATGATGCA 1344  
Db 1513 ATGAGAGAGTATGATCTCTCCGAAACCCCTCAATATACCAAGTCTTTATGATGCA 1536

RESULT 14  
AF009513 1716 bp mRNA linear ROD 01-JUN-1998  
LOCUS AF009513  
DEFINITION Mus musculus hematopoietic lineage switch 2 (HLS2) mRNA, complete cds.  
ACCESSION AF009513 GI:3169728  
VERSION AF009513.1  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1716)  
AUTHORS Williams, J.H., Chan, C.-Y. and Klinken, S.P.  
TITLE Hematopoietic lineage switch 2 (HLS2), a novel mRNA species induced during an erythroid to myeloid lineage switch  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1716)  
AUTHORS Williams, J.H., Chan, C.-Y. and Klinken, S.P.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUN-1997) Biochemistry Department (Laboratory of Cancer Medicine), University of Western Australia, Royal Perth Hospital, Wellington Street, Perth, WA 6001, Australia  
FEATURES  
source location/Qualifiers  
1. 1716  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
1. 1716  
/gene="HLS2"  
97. 1398  
/gene="HLS2"  
/codon\_start=1  
/product="hematopoietic lineage switch 2"  
/protein\_id="AAC17945.1"  
/db\_xref="GI:3169728"  
translation="MRSFLFLFVHLALSGKAVFKGVRSORTREIKELIANEYDVAKATILAVYGVKONRSYERLGLIVDVPRISGSKLEKALQIMYONLODGLFNHLEQVRIPIHMERGESAVMEBRIRHKAILIGLGSITGRPGGATIEVAVFSLQKASEARSKIIVYNDPTGYEKTVQYVGAFAVAAVAVASLIQSVAFSLSPRTGIQKYDGVKPIPTACTTVEADKMSRMAKRGKIVTHLMGAKYEDTDSFVVAITGSMYEDVQAGALDGDGAFISWEALSLVKDGLRKRTRLRLVLAETAEQGGIGASQYSELHKNISKYSLVMEADSGTFLPTGLQFTGSDPKARAIKKEVNNLLQPLNVTIKVPSNGSGTDINFWIQAVIPGASLRDDLYKYFPFPHSHSDTMTSIQSR"

BASE COUNT 476 a 370 c 414 g 456 t  
ORIGIN

Query Match 74.4%; Score 1000.4; DB 10; Length 1716;  
Best Local Similarity 85.6%; Pred. No. 2,66-273;  
Matches 1151; Conservative 0; Mismatches 156; Indels 37; Gaps 2;

Qy 1 AAGATGACATCTTTAAGAGACTTTTGAAGAAATTAAGAGAAATAGCCAGCTGTGGA 60  
Db 163 AAGATGCTGTCTCTAGCGAATCTTGAAGAAATTAAGAGAAATAGCCAGCTATGAA 222

Qy 61 GATTTGCTAAAGCAATCATCAACCTAGCTGTTATGTAAAGCCAGAACATCTTAT 120  
Db 223 GATTTGCTAAAGCAATCATCAACCTGCTGTTATGTAAATACCAAGACCGTCTAT 282  
Qy 121 GAGCGATTGGACCTTCTGTTGATTAAGTGTGAACCCAGACTGATGGCTCCAAAGACTA 180  
Db 283 GAGCGATTGGACCTTCTGTTGATTAAGTGTGAACCCAGACTGATGGCTCCAAAGACTA 342  
Qy 181 GAAAAAGCATCCAAATTATGATCAAAACCTGACGAAATGAGGCTGAGAAAGTTTAC 240  
Db 343 GAGAAAGCTATTTCAAAATCATGATCAAAACCTGACGAAATGAGGCTGAGAAAGTTTAC 402  
Qy 241 CTGAGCGATGAGAAATATACCCCATGAGGAGAGGAGAAAGATCAGCTGATGCTGAG 300  
Db 403 CTGAGCGATGAGAAATATACCCCATGAGGAGAGGAGAAAGATCAGCTGATGCTGAG 462  
Qy 301 CCAAGAAATTCATTAAGATAGCCATCTGAGTCTTGGACAGCATTTGGACTCTCCAGAA 360  
Db 463 CTGAGATTCACAAAGATGCTATTTCTGATGCTTGGACAGCATTTGGACTCTCCAGAA 522  
Qy 361 GGCATTCACAGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
Db 523 GGCATTCACAGAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 582  
Qy 421 GAAGCAAGAGGAAAGATTTGTTTATTAACCAACCTTACATCACTCAAGGACGGTG 480  
Db 583 GAAGCAAGAGGAAAGATTTGTTTATTAACCAACCTTACATCACTCAAGGACGGTG 642  
Qy 481 CAATACCGAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
Db 643 CAATACCGAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 702  
Qy 541 CGATCGGAGGCTCTCTTCTCATGATGATGATGATGATGATGATGATGATGATGATGAT 600  
Db 703 CAATACGATGATCTCTCTTCTCATGATGATGATGATGATGATGATGATGATGATGATGAT 762  
Qy 601 GAGCGTCCCAAAATTCACAGGCTGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660  
Db 763 GAGCGTCCCAAAATTCACAGGCTGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 822  
Qy 661 ATGCGTCTCATGAGGATCAAAATTTGTCATTCAGCTAAAGATGAGGAGGAGGAGGAGGAGG 720  
Db 823 ATGCGTCTCATGAGGATCAAAATTTGTCATTCAGCTAAAGATGAGGAGGAGGAGGAGGAGG 882  
Qy 721 GATTCGATTCCTTCAACACCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
Db 883 GATTCGATTCCTTCAACACCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 935  
Qy 781 GTACTGTCAGTGGACATCTGACAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840  
Db 936 -----GAGTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 966  
Qy 841 GGTGAGACCTTTATATCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900  
Db 967 GGTGAGACCTTTATATCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1026  
Qy 901 AAGAGACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
Db 1027 AAGAGACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1086  
Qy 961 CAGATTTATAGTTACACAGGATTAATTTTCAACATCAAGTCTGATGATGATGATGATGAT 1020  
Db 1087 CAGATTTATAGTTACACAGGATTAATTTTCAACATCAAGTCTGATGATGATGATGATGATGAT 1146  
Qy 1021 GCAAGAACTTTTATACCACTGGGCTGCAATTCATGCTGCAAGTGAAGGAGGAGGAGGAGGAGG 1080  
Db 1147 TCAGGAACCTTTTATACCACTGGGCTGCAATTCATGCTGCAAGTGAAGGAGGAGGAGGAGGAGG 1206  
Qy 1081 ATGAGAGAGGTTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
Db 1207 ATGAGAGAGGTTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1266



QY 1141 GAAGGACAGACATCACTTTTGGATCCAGCTGGAGTGCCTGGAGCCAGCTCTACTGAT 1200  
DB 1267 GAAGGACAGCTGATTAATCTTCTGATCCAGCTGAGTACCTGAGGAGCTGCTGAT 1326  
QY 1201 GACTTATACAAAGATTTCTTCTTCATCACTCCACGAGACACCATGACTGTCAATGAT 1260  
DB 1327 GACTGTACAGATTTCTTCTTCATCACTCCACGAGACACCATGAC-GTCATGAT 1385  
QY 1261 CCAAGAGAGATGATGTCGTCGCTGTTGGGCGTGTGTTCTTATGTTGTGCGAGAC 1320  
DB 1386 CCAAGAGAGATGATGTCGTCGCTGTCGTATGAGGCTGTGTCGCTTATGTTGTGCGAGAC 1445  
QY 1321 ATGAGAGAAATGCTGCTAGCTCC 1344  
DB 1446 ATGATGAATGCTGCCAGCTCC 1469

RESULT 15  
BD059610 609 bp DNA linear PAT 27-AUG-2002  
LOCUS Secreted expressed sequence tags (ESTs).  
DEFINITION BD059610  
ACCESSION BD059610.1 GI:22605216  
VERSION JP 2001519666-A/1465.  
KEYWORDS Zea mays  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
Clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 609)  
Jacob, K., McCoy, J. M., Lavallie, E. R., Racie, L. A., Merberg, D.,  
Treacy, M., Spaulding, V. and Agostino, M. J.  
Secreted expressed sequence tags (ESTs).  
Patent: JP 2001519666-A 1465 23-OCT-2001;  
GENETICS INSTITUTE INC  
FN JP 2001519666-A/1465  
PD 23-OCT-2001  
RF 10-APR-1998 JP 1998543068  
FR 10-APR-1997 US 08/835913  
FI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI  
DAVID MERBERG,  
PI MAURICE TREACY, VIRKI SPAULDING, MICHAEL J AGOSTINO PC  
C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:  
Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.  
FEATURES  
Source Location/Qualifiers  
1..609  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"  
BASE COUNT 190 a 127 c 148 g 144 t  
ORIGIN

Query Match 33.4%; Score 448.6; DB 6; Length 609;  
Best Local Similarity 98.9%; Pred. No. 3.7e-116;  
Matches 462; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 AAGATGGCATCTCTAAGAGACTTTTGAAGAAATTAAGAAAGAAATAGCCAGCTGTGA 60  
DB 143 AAGATGGCATCTCTAAGAGACTTTTGAAGAAATTAAGAAAGAAATAGCCAGCTGTGA 202  
QY 61 GATTTTGTAAAGCAATCATCACTAGCTGTTTATGTAAAGCCAGAAACAGATCTTAT 120  
DB 203 GATTTTGTAAAGCAATCATCACTAGCTGTTTATGTAAAGCCAGAAACAGATCTTAT 262  
QY 121 GAGGATTTGGCACTTCTGTTGATAGTGTGACCCAGACTGAGTGCTCCAGAACTTA 180  
DB 263 GAGGATTTGGCACTTCTGTTGATAGTGTGACCCAGACTGAGTGCTCCAGAA-CTA 321  
QY 181 GAAAGAGCCATCAAAATTATGTACMAAACTGAGCAAGATGGGCTGAGAAAGTTTAC 240  
DB 322 GAAAGAGCCATCAAAATTATGTACMAAACTGAGCAAGATGGGCTGAGAAAGTTTAC 381

QY 241 CTGAGCCAGTGAATATACCCCACTGGAGAGGGGAGAAAGATTCAGCTGTGATCTGGAG 300  
DB 382 CTGAGCCAGTGAATATACCCCACTGGAGAGGGGAGAAAGATTCAGCTGTGATCTGGAG 441  
QY 301 CCAAGATTCATTAAGATAGCATCTGCTGCTTGGAGAGCAATGGGACTCTCCAGAA 360  
DB 442 CCAAGATTCATTAAGATAGCATCTGCTGCTTGGAGAGCAATGGGACTCTCCAGAA 501  
QY 361 GGCATTACAGCAAGATTTCTGTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 420  
DB 502 GGCATTACAGCAAGATTTCTGTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 561  
QY 421 GAAGCAAGAGGAAATTTGTTTATTAACCAACTTACATCACTA 467  
DB 562 GAAGCAAGAGGAAATTTGTTTATTAACCAACTTACATCACTA 608

Search completed: December 22, 2003, 20:22:22  
Job time : 4856.44 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 12:51:18 ; Search time 323.93 Seconds

(without alignments)  
11200.072 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_171\_1514

Perfect score: 1344

Sequence: 1 AAGAAATGCATCTCTTAAGAG.....AAGAAATGCTCCTAGCTCC 1344

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1344	100.0	1851	19	AAV82779
2	1344	100.0	1851	24	ABO92016
3	1340.8	99.8	1767	21	AAA40493
4	1340.8	99.8	1778	18	AAV02296
5	1340.8	99.8	1863	21	AAZ96034
6	1340.8	99.8	1863	22	AAD11647
7	1340.8	99.8	1863	24	ABK69743
8	1340.8	99.8	1863	25	ACC50817

9	1340.8	99.8	1863	25	AB271453
10	1340.8	99.8	1884	21	AA258313
11	1340.8	99.8	1923	21	AA298139
12	1340.8	99.8	2077	25	ACC50603
13	1340.8	99.8	2077	25	AB271331
14	1336	99.4	1860	22	AAK94491
15	1283.6	95.5	1895	22	AAH99703
16	1224.8	91.1	1784	23	AA573592
17	1060.8	78.9	1778	24	ABK62598
18	772.4	57.5	1134	21	AA298101
19	772.4	57.5	1134	22	AA011714
20	772.4	57.5	1134	24	ABK69810
21	772.4	57.5	1134	25	ACC50818
22	772.4	57.5	1134	25	AB271454
23	770.8	57.4	895	21	AAA44369
24	448.6	33.4	609	20	AAV87487
25	398.2	29.6	642	22	AAK92181
26	398.2	29.6	642	22	AAK93447
27	212.8	15.8	314	20	AAV86369
28	204.6	15.2	462	18	AA797398
29	204.6	15.2	462	18	AA788060
30	204.6	15.2	462	18	AAV02319
31	170.4	12.7	441	22	AAK56831
32	140.6	10.5	317	20	AAK40587
33	114.2	8.5	492	22	AAK92932
34	90.8	6.8	394	14	AA061260
35	61.4	4.6	424	23	AA573591
36	60.2	4.5	65	24	ABN53183
37	60	4.5	60	24	ABN36212
38	42	3.1	784	22	AAH05499
39	42	3.1	2200	22	AAH15906
40	42	3.1	2200	22	AAH78065
41	42	3.1	2408	22	AAH99550
42	41	3.1	65	24	ABN28454
43	39.4	2.9	975	21	AACT7713
44	39.4	2.9	1556	24	AA033726
45	39.4	2.9	2319	25	ABX63579

## ALIGNMENTS

RESULT 1  
AAV82779  
ID AAV82779 standard; cDNA; 1851 BP.  
XX  
AC AAV82779;  
XX  
DT 25-FEB-1999 (first entry)  
XX  
DB Clone bu45\_2 isolated from human adult placenta cDNA library.  
XX  
KW Secreted protein; nutritional activity; immune stimulating; vaccine;  
KW \* suppressing activity; haematopoiesis regulating activity;  
KW tissue growth activity; activin; inhibin activity; chemotaxis;  
KW chemokine activity; haemostasis; thrombolytic activity; receptor;  
KW ligand; anti-inflammatory; cadherin; tumour invasion suppressor;  
KW tumour inhibition; gene therapy; ds.  
XX  
OS Homo sapiens.  
XX  
PN W09842739-A2.  
XX  
PD 01-OCT-1998.  
XX  
PF 20-MAR-1998; 98WO-US05653.  
XX  
PR 19-MAR-1998; 98US-0044466.  
XX  
PR 21-MAR-1997; 97US-0822167.  
XX  
PA (GENY) GENETICS INST INC.  
XX  
PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

PI Racine LA, Spaulding V, Treacy M;  
 XX MPI, 1998-609890/51.  
 DR P-F8DB; AAM85456.  
 XX  
 PT New polynucleotides encoding secreted human proteins - derived from  
 PT human foetal brain, adult brain, foetal kidney, placenta or adult  
 PT pineal gland cDNA libraries.  
 XX  
 PS Claim 14; Page 69-70; 113pp; English.  
 XX  
 CC The present sequence encodes a secreted protein. The polynucleotide and  
 CC secreted protein are predicted to have biological activities which would  
 CC make them suitable for treating, preventing or ameliorating medical  
 CC conditions in humans and animals, although no supporting data is given.  
 CC Suggested activities include nutritional activity, immune stimulating  
 CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating  
 CC activity, tissue growth activity, activin/inhibin activity,  
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,  
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour  
 CC invasion suppressor activity, and tumour inhibition activity (no data is  
 CC given in the specification to support these activities). The  
 CC polynucleotide is also stated to be useful for gene therapy.  
 CC  
 SQ Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;  
 Query Match 100.0%; Score 1344; DB 19; Length 1851;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGATGCGATCTCTAAGAGACCTTTGAGAAATGAAAGAAATGAGCCAGCTGTGGA 60  
 DB 171 AAGATGCGATCTCTAAGAGACCTTTGAGAAATGAAAGAAATGAGCCAGCTGTGGA 230  
 QY 61 GATGTGCTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCCAAGACGATCTAT 120  
 DB 221 GATGTGCTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCCAAGACGATCTAT 290  
 QY 121 GAGCGATTGGCACTCTGGTTGATCTGTGGAACCAAGCTGAGTGGCTCCAGAAACCTA 180  
 DB 291 GAGCGATTGGCACTCTGGTTGATCTGTGGAACCAAGCTGAGTGGCTCCAGAAACCTA 350  
 QY 181 GAAAAAGCATCAATTTATGTACAAAACTGACAGAAATGAGCTGAGAAATGTTAC 240  
 DB 351 GAAAAAGCATCAATTTATGTACAAAACTGACAGAAATGAGCTGAGAAATGTTAC 410  
 QY 241 CTGAGCCAGTAGAATATCCCACTGGAGAGAGGGAGAAATCACTGTGTGCTGAG 300  
 DB 411 CTGAGCCAGTAGAATATCCCACTGGAGAGAGGGAGAAATCACTGTGTGCTGAG 470  
 QY 301 CCAAGAATTCAATAGATGCGATCTGGTCTTGGCAGAGATTTGGATCTCTCAGAA 360  
 DB 471 CCAAGAATTCAATAGATGCGATCTGGTCTTGGCAGAGATTTGGATCTCTCAGAA 530  
 QY 361 GGCATTACAGCAGAGATTCTGGTGTGACCTCTTTCAGATGACCTGAGAGAGGCTCA 420  
 DB 531 GGCATTACAGCAGAGATTCTGGTGTGACCTCTTTCAGATGACCTGAGAGAGGCTCA 590  
 QY 421 GAAAGCAAGAGGAGATTGTTGTTAAACCAACTTATCATCACTACTCAAGAGCGGTG 480  
 DB 591 GAAAGCAAGAGGAGATTGTTGTTAAACCAACTTATCATCACTACTCAAGAGCGGTG 650  
 QY 481 CAATCCCAAGCGAGGGCGGTGGAAGCTGCCAAGTGGGGCTTTGGCACTCTCTATT 540  
 DB 651 CAATCCCAAGCGAGGGCGGTGGAAGCTGCCAAGTGGGGCTTTGGCACTCTCTATT 710  
 QY 541 CGATCCGTGGCTCTCTTCATCTACAGTCTCTCAACAGATTCAGAGATACAGAGAT 600  
 DB 711 CGATCCGTGGCTCTCTTCATCTACAGTCTCTCAACAGATTCAGAGATACAGAGAT 770  
 QY 601 GGCCTGCCCAAAATTCACACGCTGTATTCAGGTGAGAAATGATGTCAAGA 660  
 DB 771 GGCCTGCCCAAAATTCACACGCTGTATTCAGGTGAGAAATGATGTCAAGA 830

QY 661 ATGGCTTCTCATGGGATCAAAATTTGTCATTGCTAATAAGATGGGGCAAGACTTACCA 720  
 DB 831 ATGGCTTCTCATGGGATCAAAATTTGTCATTGCTAATAAGATGGGGCAAGACTTACCA 890  
 QY 721 GATACGATATCTTCAACACTGTGAGAGATCACTGGAGCAAAATTCAGAACAGGTT 780  
 DB 891 GATACGATATCTTCAACACTGTGAGAGATCACTGGAGCAAAATTCAGAACAGGTT 950  
 QY 781 GTACTGTGCTGAGACATCTGAGACAGCTGGAGATCTTGGAGAGGTTGCATGATGATGC 840  
 DB 951 GTACTGTGCTGAGACATCTGAGACAGCTGGAGATCTTGGAGAGGTTGCATGATGATGC 1010  
 QY 841 GGTGAGGCTTTATATATGAGGAGCACTCTCACTTATTAAGATCTTGGGCTCGTCA 900  
 DB 1011 GGTGAGGCTTTATATATGAGGAGCACTCTCACTTATTAAGATCTTGGGCTCGTCA 1070  
 QY 901 AAGAGACTCTGGCGCTGTGCTGTGACCTGACAGAGAAACAAGTGAAGTTGCTTC 960  
 DB 1071 AAGAGACTCTGGCGCTGTGCTGTGACCTGACAGAGAAACAAGTGAAGTTGCTTC 1130  
 QY 961 CAGTATTATCAGTTACCAAGGTAAATTTTCCATCACTGCTGTGATGATGATCTGAC 1020  
 DB 1131 CAGTATTATCAGTTACCAAGGTAAATTTTCCATCACTGCTGTGATGATGATCTGAC 1190  
 QY 1021 GCAGAACTCTTACCACTGGGCTGCAATTCCTGAGAGTGAAGGCAAGGCGCATC 1080  
 DB 1191 GCAGAACTCTTACCACTGGGCTGCAATTCCTGAGAGTGAAGGCAAGGCGCATC 1250  
 QY 1081 ATGAGAGAGGTTATGAGCTGTGACAGCCCTCAATATCACTGAGTCTGAGCCATGA 1140  
 DB 1251 ATGAGAGAGGTTATGAGCTGTGACAGCCCTCAATATCACTGAGTCTGAGCCATGA 1310  
 QY 1141 GAAGGACAGACATCACTTTGATCCAACTGAGATGCTCTGAGGCAAGTCACTGAT 1200  
 DB 1311 GAAGGACAGACATCACTTTGATCCAACTGAGATGCTCTGAGGCAAGTCACTGAT 1370  
 QY 1201 GACTTATACAGATATTTCTTCTTCATCACTCCCAAGAGACATGATCATGAT 1260  
 DB 1371 GACTTATACAGATATTTCTTCTTCATCACTCCCAAGAGACATGATCATGAT 1430  
 QY 1261 CCAAGCAGATGAATGTTGCTGCTGTGTTGGCTGTGTTCTTATGTTTGCAGAC 1320  
 DB 1431 CCAAGCAGATGAATGTTGCTGCTGTGTTGGCTGTGTTCTTATGTTTGCAGAC 1490  
 QY 1321 ATGGAAGAAATGCTGCTTAAAGTCC 1344  
 DB 1491 ATGGAAGAAATGCTGCTTAAAGTCC 1514

RESULT 2  
 ABO92016  
 ID ABO92016 standard; cDNA; 1851 BP.  
 XX  
 AC ABO92016;  
 XX  
 DT 04-OCT-2002 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 13.  
 XX  
 KW Human; cytosolic; antirheumatic; antiarthritic; vulnery; analgesic;  
 KW antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;  
 KW neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;  
 KW antitumor; fungicide; antidiabetic; antiaschemic; antiallergic;  
 KW immunostimulant; antiparasitic; secreted protein; transmembrane protein;  
 KW cytokine; cell proliferation; cell differentiation; autoimmune disease;  
 KW stem cell; growth factor; nervous system disease; neuropathy;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;  
 KW osteoporosis; severe combined immunodeficiency; SCID; infection;  
 KW multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX

EN US2002065394-A1.  
XX 30-MAY-2002.  
XX 22-DEC-2000; 2000US-0745763.  
XX 18-MAR-1998; 98US-0040963.  
XX (JACO/) JACOBS K.  
XX (MCCO/) MCCOY J M.  
XX (LAVA/) LAVALLIE E R.  
XX (COLL/) COLLINS-RACIE L A.  
XX (EVAN/) EVANS C.  
XX (MERB/) MERBERG D.  
XX (TREA/) TREACY M.  
XX (SPAU/) SPAUDING V.  
XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
XX Merberg D, Treacy M, Spaulding V;  
XX WPI; 2002-582343/62.  
XX P-PSDB; ABP61800.  
XX Novel secreted or transmembrane protein and polynucleotide encoding the  
XX PT protein, useful for diagnosis and treatment of neurological disorders,  
XX PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis  
XX  
XX Claim 50; Page 113-114; 284pp; English.  
XX  
XX The invention relates to human secreted or transmembrane protein (I),  
XX their fragments and is encoded by specific complementary deoxyribonucleic  
XX acid (CDNA) inserts (II), where the protein is substantially free from  
XX other mammalian proteins. (I) are useful for preventing, treating or  
XX ameliorating a medical condition, especially immunological treatment or  
XX prevention of tumours. (I) exhibits activity relating to angiogenesis,  
XX cytokine, cell proliferation, cell differentiation, anti-inflammatory,  
XX stem cell growth factor activity and activin or inhibin-related  
XX activities. (I) can be used to manipulate stem cells in culture to give  
XX rise to neuroepithelial cells that can be used to augment or replace  
XX cells damaged by illness, autoimmune disease, accidental damage or  
XX genetic disorders. (I) induces the proliferation of neural cells and  
XX regeneration of nerve and brain tissue and is useful for the treatment of  
XX central and peripheral nervous system diseases and neuropathies, such as  
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX lateral sclerosis. (I) is involved in chemotactic or chemokinetic  
XX activity, regulation of haematopoiesis and is useful for treating myeloid  
XX or lymphoid cell disorders, platelet disorders such as thrombocytopenia  
XX and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
XX tissue growth and in tissue repair, healing of burns, incisions, ulcers,  
XX for treating osteoporosis, osteoarthritis, bone degenerative disorders or  
XX periodontal disease. (I) is also useful for gut protection or  
XX regeneration and treatment of lung or liver fibrosis, reperfusion injury  
XX in various tissues, various immune deficiencies and disorders including  
XX severe combined immunodeficiency (SCID), bacterial or fungal infections,  
XX autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,  
XX diabetes mellitus, myasthenia gravis, allergic reactions and conditions,  
XX such as asthma or other respiratory problems. (II) is useful to express  
XX recombinant protein, as markers for tissues in which the corresponding  
XX protein is preferentially expressed and in gene therapy. The present  
XX sequence is that of a polynucleotide of the invention.  
XX  
XX Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;  
XX  
XX Query Match 100.0%; Score 1344; DB 24; Length 1851;  
XX Best Local Similarity 100.0%; Pred. No. 0;  
XX Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 231 GATGTTGCTAAAGCAATCATCAACTAGCTGTTATGTAAAGCCAGAAAGATCTTAT 290  
Qy 121 GAGCGATTGGGACCTTCGTTGATGATCTGTTGACCCGACGCTGCTCCAGAACCTTA 180  
Db 291 GAGCGATTGGGACCTTCGTTGATGATCTGTTGACCCGACGCTGCTCCAGAACCTTA 350  
Qy 181 GAAAGAGCCATCCAAATTAATGTAACCAAACTGACGACGAGATGGGCTGGAGAAATTCAC 240  
Db 351 GAAAGAGCCATCCAAATTAATGTAACCAAACTGACGACGAGATGGGCTGGAGAAATTCAC 410  
Qy 241 CTGGAGCCAGTGAATATCCCACTGGAGAGGGGAGAGAAATCAGCTGTGATCTGGAG 300  
Db 411 CTGGAGCCAGTGAATATCCCACTGGAGAGGGGAGAGAAATCAGCTGTGATCTGGAG 470  
Qy 301 CCAGAAATTCATAGATAGGCAATCCTGGGCTTGGGACGAGCAATGGGACTCCCTCCAGAA 360  
Db 471 CCAGAAATTCATAGATAGGCAATCCTGGGCTTGGGACGAGCAATGGGACTCCCTCCAGAA 530  
Qy 361 GGCATTACAGCAGAGATTTGTTGTTGATGACCTTTGATGATGACGAGAGAGGCTTCA 420  
Db 531 GGCATTACAGCAGAGATTTGTTGTTGATGACCTTTGATGATGACGAGAGAGGCTTCA 590  
Qy 421 GAAGCAAGAGGAAAGATTTGTTGTTGATGATGACCTTTGATGATGACGAGAGAGGCTTCA 480  
Db 591 GAAGCAAGAGGAAAGATTTGTTGTTGATGATGACCTTTGATGATGACGAGAGAGGCTTCA 650  
Qy 481 CAATACCGAAGCAGAGGGGGGGGAGAGAGCTGCAAGAGTGGGGCTTGGATCTTCAT 540  
Db 651 CAATACCGAAGCAGAGGGGGGGGAGAGAGCTGCAAGAGTGGGGCTTGGATCTTCAT 710  
Qy 541 CGATCCGTCGCTCTCTCTCATCTACAGTCTCAACAGATTTGATGATGATGATGATGAT 600  
Db 711 CGATCCGTCGCTCTCTCTCATCTACAGTCTCAACAGATTTGATGATGATGATGATGAT 770  
Qy 601 GGGCTGCCCCAAATTTCCAAAGCTGTATTACGGTGAAGATGCAAAATGATGATGATGAT 660  
Db 771 GGGCTGCCCCAAATTTCCAAAGCTGTATTACGGTGAAGATGCAAAATGATGATGATGAT 830  
Qy 661 ATGGCTTCTATGGAGATCAAAATTTGATGATGATGATGATGATGATGATGATGATGAT 720  
Db 831 ATGGCTTCTATGGAGATCAAAATTTGATGATGATGATGATGATGATGATGATGATGAT 890  
Qy 721 GATCTGATTCCTTCAACAGTGTAGAGATCACTGGAGCAATATTCAGAACAGATT 780  
Db 891 GATCTGATTCCTTCAACAGTGTAGAGATCACTGGAGCAATATTCAGAACAGATT 950  
Qy 781 GTACTGTCAAGTGAACATCTGACAGCTGGAGATTTGGGACAGGCTGCATGATGATGATGAT 840  
Db 951 GTACTGTCAAGTGAACATCTGACAGCTGGAGATTTGGGACAGGCTGCATGATGATGATGAT 1010  
Qy 841 GGTGAGACCTTTATATATGAGGAAGCACTCACTTTTAAGATCTTTGGCTGCTGCCA 900  
Db 1011 GGTGAGACCTTTATATATGAGGAAGCACTCACTTTTAAGATCTTTGGCTGCTGCCA 1070  
Qy 901 AAGAGACCTGCGGGCTGGTGTCTGAGTCAAGAAAGCAAGGTGAGTGGTGGCTTC 960  
Db 1071 AAGAGACCTGCGGGCTGGTGTCTGAGTCAAGAAAGCAAGGTGAGTGGTGGCTTC 1130  
Qy 961 CAGTATTATCAGTTACCAAGATTAATTTTCCAACTCAAGTCTGTGATGATGATGATGATGAT 1020  
Db 1131 CAGTATTATCAGTTACCAAGATTAATTTTCCAACTCAAGTCTGTGATGATGATGATGATGAT 1190  
Qy 1021 GCAGGAACCTTTTACCACTGCGGCTGCATTCCTGCGAGTGAAGGCGGAGGCTATC 1080  
Db 1191 GCAGGAACCTTTTACCACTGCGGCTGCATTCCTGCGAGTGAAGGCGGAGGCTATC 1250  
Qy 1081 ATGAGAGAGGTTATGAGCTGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 1140  
Db 1251 ATGAGAGAGGTTATGAGCTGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 1310  
Qy 1141 GAAAGGACAGACATCACTTTTGGATCAAGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAT 1200

```

Db      1311  GAAGGACAGACATCAACTTTTGGATFCCAAGTGGAGTGGCCAGCTCACTTGTAT 1370
Oy      1201  GACTTATATCAAGTATTTCTTCTTCATCACTCTCCACGGAGACATGATCTCATGGAT 1260
Db      1371  GACTTATATCAAGTATTTCTTCTTCATCACTCTCCACGGAGACATGATCTCATGGAT 1430
Oy      1261  CCAAGACGATGATGATGTTGCTGCTGCTGTTTGGGGCTGTTTGTCTTATGTTGTGCAGAC 1320
Db      1431  CCAAGACGATGATGATGTTGCTGCTGCTGTTTGGGGCTGTTTGTCTTATGTTGTGCAGAC 1490
Oy      1321  ATGGAAGAAATGCTGCTCCTAGGTCC 1344
Db      1491  ATGGAAGAAATGCTGCTCCTAGGTCC 1514

RESULT 3
AAAA0493
ID      AAAA0493 standard; cDNA; 1767 BP.
XX
AC      AAAA0493;
XX
DT      16-NOV-2000 (first entry)
XX
DE      Human fetal kidney cDNA fragment AM282_11.
XX
KW      Secreted protein; cytosolic; immunostimulatory; antimicrobial;
KW      antiviral; immunosuppressive; antiinflammatory; vulnery; cytokine;
KW      cell proliferation; differentiation; regulator; treatment; tumor;
KW      autoimmune disease; inflammatory disorder; wound; microbial infection;
KW      viral disease; graft versus host reaction suppression; ss.
XX
XX      Homo sapiens.
OS
XX      WO200037630-A1.
PN
XX
PD      29-JUN-2000.
XX
PF      22-DEC-1999; 99WO-US31005.
XX
PR      23-DEC-1998; 98US-0220876.
XX
PA      (GENY ) GENETICS INST INC.
PI      Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
PI      Metberg D, Treacy M, Bowman MR;
XX
DR      WPI; 2000-442661/38.
DR      P-PSDB; AAB10229.
XX
FT      Secreted human proteins AS296-11 and AS34-11, useful for treating
FT      tumors, autoimmune diseases, inflammatory disorders, wounds, microbial
FT      infections and viral diseases -
XX
PS      Disclosure, Page 198, 293pp; English.
XX
XX      This invention describes novel secreted human proteins (I) which have
XX      cytosolic, immunostimulatory, antimicrobial, antiviral,
XX      immunosuppressive, antiinflammatory and vulnery activity and which act
XX      as cytokine, cell proliferation or differentiation regulators. (I)
XX      is useful for treating tumors, autoimmune diseases, inflammatory
XX      disorders, wounds, microbial infections and viral diseases. (I) is also
XX      useful for suppressing graft versus host reaction. AAA40490-A40580
XX      represent cDNA fragments that encode the secreted proteins
XX      AAB10226-B10288 described in the method of the invention.
SQ      Sequence 1767 BP; 512 A; 384 C; 418 G; 453 T; 0 other;

Query Match      99.8%; Score 1340.8; DB 21; Length 1767;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AAGATGGCATCTCTAAGAGGACTTTGAAGAAATTAAGAAAGAAATAGCCAGCTGTGGA 60
|||||

```

Db	78	AAGATGGCATCTCTAAGAGGACTTTTGAAGAAATPAAAAAGAAATAGCCAGCTGTGG	137
QY	61	GATGTTGCTTAAGCAATCATCACTAGCTGTTATGTGTAAGCCAGAACAGATCTAT	120
Db	138	GATGTGCTTAAGCAATCATCACTAGCTGTTATGTGTAAGCCAGAACAGATCTAT	197
QY	121	GAGGAGTTGGACCTTCGTGGTGTGATACGTGTGGACCCACAGAGTGGCTCCAGAAACCTA	180
Db	198	GAGGAGTTGGACCTTCGTGGTGTGATACGTGTGGACCCACAGAGTGGCTCCAGAAACCTA	257
QY	181	GAAGAAAGCCATCCAAATTTATGTACCAAAACCTGACAGCAAGATGGCTGAGAAATTTAC	240
Db	258	GAAGAAAGCCATCCAAATTTATGTACCAAAACCTGACAGCAAGATGGCTGAGAAATTTAC	317
QY	241	CTGAGACCACTGAGAAATACCCCACTGGGAGAGGGAGAAATACGCTGTGATGCTGGAG	300
Db	318	CTGAGACCACTGAGAAATACCCCACTGGGAGAGGGAGAAATACGCTGTGATGCTGGAG	377
QY	301	CCAAGAAATTCATTAAGATAGCCATCTGGGATCTTGGACACATTTGGGACTCTCCAGAA	360
Db	378	CCAAGAAATTCATTAAGATAGCCATCTGGGATCTTGGACACATTTGGGACTCTCCAGAA	437
QY	361	GGCATTTACAGCAGAACTTCGTGGTGTGACCTCTTTGCATGCACTGACAGAAAGGGCTCA	420
Db	438	GGCATTTACAGCAGAAAGTTCTGTGTGTGACCTCTTTGCATGCACTGACAGAAAGGGCTCA	497
QY	421	GAAGCAGAGGGAAAGATTTGTTTATTAACCACTTACATCACTACAGAGCGGTG	480
Db	498	GAAGCAGAGGGAAAGATTTGTTTATTAACCACTTACATCACTACAGAGCGGTG	557
QY	481	CAATACCGAAGAGGAGGGGGCGGTGGAAAGCTCCAAAGTGGGGGCTTGGACATCTCATTT	540
Db	558	CAATACCGAAGAGGAGGGGGCGGTGGAAAGCTCCAAAGTGGGGGCTTGGACATCTCATTT	617
QY	541	CGATCCGGGCGCTCCTTCTCCATCTACAGTCTCTACAAGGTATTTCAAGAAATCACAGAT	600
Db	618	CGATCCGGGCGCTCCTTCTCCATCTACAGTCTCTACAAGGTATTTCAAGAAATCACAGAT	677
QY	601	GGCGGTGCCAAATAATTCCAACAGCCTGTATTACGGTGGAAAGATGCAGAAATGATGTCAAGA	660
Db	678	GGCGGTGCCAAATAATTCCAACAGCCTGTATTACGGTGGAAAGATGCAGAAATGATGTCAAGA	737
QY	661	ATGGCTTCTCATGGGATTCAAAATTGTCAATTAGCTTAAAGATGGGGGGCAAAGACTTACCCA	720
Db	738	ATGGCTTCTCATGGGATTCAAAATTGTCAATTAGCTTAAAGATGGGGGGCAAAGACTTACCCA	797
QY	721	GATACGTGATTCCTTCAACACTGTAGCAGAGATTCACCTGGGAGCAAAATCCAGAACAGGTT	780
Db	798	GATACGTGATTCCTTCAACACTGTAGCAGAGATTCACCTGGGAGCAAAATTCAGAACAGGTT	857
QY	781	GTACTGTGCTAGTGACATCTGAGCAGCTGGGATGTTGGGAGGGGTGCCATGATGATGGC	840
Db	858	GTACTGTGCTAGTGACATCTGAGCAGCTGGGATGTTGGGAGGGGTGCCATGATGATGGC	917
QY	841	GGTGGAGCTTTAATCATGGGAAGCACTCTCATTTTAAAGATCTTGGGCTGGCTCCA	900
Db	918	GGTGGAGCTTTAATCATGGGAAGCACTCTCATTTTAAAGATCTTGGGCTGGCTCCA	977
QY	901	AAAGAGGACTCTGGGGCTGGTCTGAGACCTGACAGAAAGAAAGAGGTGGAGTGGCTTTC	960
Db	978	AAAGAGGACTCTGGGGCTGGTCTGAGACCTGAGACAGAGAAAGAGGTGGAGTGGCTTTC	1033
QY	961	CAGTATTTACGTTTACCAAGAGTAAATATTTCCAATCAAGTCTGTGATGGAAGCTTGAC	1022
Db	1038	CAGTATTTACGTTTACCAAGAGTAAATATTTCCAATCAAGTCTGTGATGGAAGCTTGAC	1099
QY	1021	GCAGAGAACTTCTTAACTCACTGGCTGCAATTCACCTGGCAGTGAAGAAAGCCAGGGCATC	1086
Db	1098	GCAGAGAACTTCTTAACTCACTGGCTGCAATTCACCTGGCAGTGAAGAAAGCCAGGGCATC	1157
QY	1081	ATGAGAGAGGTTATGAGCGCTGCTGACAGCCCTCAATATATCACTCAAGTCTGAGGCATGGA	1146
Db	1158	ATGAGAGAGGTTATGAGCGCTGCTGACAGCCCTCAATATATCACTCAAGTCTGAGGCATGGA	1217

Oy		1141	GAGGGACAGCATTCACTTTTGGATCCAAAGCTGGAGTCCTGAGCCAGTCACTTGAT	1207
Dd		1218	GAGGGACAGCATCACTTAACCTTTTGGATCCAAGCTGGAAGTCCCGAAGCAAGTCACTTGAT	1277
Oy		1301	GACTTATACAAGTATTTCTTCTTCATCACTCCACGAGAGACACCATTGACTGTATGGAT	1266
Dd		1278	GACTTATACAAGTATTTCTTCTTCATCACTCCACGAGAGACACCATTGACTGTATGGAT	1337
Oy		1261	CCAAAAGCAGATGATGTTGCTGCTGCTGTTGGAGCTGTTGTTTCTTATGTTGTCAGAC	1320
Dd		1338	CCAAAAGCAGATGATGTTGCTGCTGCTGTTGGAGCTGTTGTTTCTTATGTTGTCAGAC	1397
Oy		1321	ATGGAAGAATACTGCTCCTTAGTCC 1344	
Dd		1398	ATGGAAGAATACTGCTCCTTAGTCC 1421	
 RESULT 4 AAV02296 AAV02296 standard; DNA; 1778 BP.				
XX	AAV02296;			
AC	AAV02296;			
XX				
DT	21-MAY-1998	(first entry)		
XX				
DE	Human secreted protein AM282 full-length cDNA clone.			
KX	Secreted protein; AM282; cytokine; human; ds.			
XX				
OS	Homo sapiens.			
XX				
FT	Key	Location/Qualifiers		
FH	CDS	/tag= a		
FT		17..1435		
FT	sig_peptide	/tag= b		
FT		17..88		
FT	mat_peptide	/tag= b		
FT		89..1432		
XX		/tag= b		
PN	W09739030-A2.			
XX				
PD	23-OCT-1997.			
XX				
PF	16-APR-1997;	97WO-US06475.		
PR	13-JAN-1997;	97US-0783520.		
PR	18-APR-1996;	96US-0634325.		
XX				
PA	(GENEY ) GENETICS INST INC.			
PI	Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA;			
PI	Sputating V;			
XX				
DR	WPI; 1997-526400/48.			
XX	P-PSDB; AAAM33604.			
PT	New isolated secretory proteins AM340, AM282 and AK583 - possibly have cytokine, cell proliferation/differentiation regulating,			
PT	Immunomodulating activities, etc.			
XX				
PS	Claim 15; Page 44-45; 59pp; English.			
XX				
CC	This cDNA clone encodes a protein (see W33604) designated AM282.			
CC	It was identified as "yfg5b10.r1 human EST 30142 5" (Genbank accession No. R77830) in a database search using a partial AM282 clone (see T97398) obtained from a human foetal kidney cDNA library using methods selective for cDNAs encoding secreted proteins. AM282 is deposited in ATCC 98026 together with clones AM340 (see T97397) and AK583 (see V02297), which are also claimed.			
CC	AM282 protein can be used in a claimed method for preventing, treating or ameliorating a medical condition. It may exhibit cytokine, cell proliferation (either inducing or inhibiting) or			

	Query Match	99.8%	Score 1340.8	DB 18	Length 1778	
	Best Local Similarity	99.9%	Pred. No. of			
	Matches 1342	Conservative 0	Mismatches 2	Indels 0	Gaps 0	
CC	cell differentiation (either inducing or inhibiting) activity or					
CC	may induce production of other cytokines in certain cell					
CC	populations. It may also exhibit e.g. immune stimulating or					
CC	suppressing activity, haematopoiesis regulating activity, tissue					
CC	growth activity, activin/inhibin activity, chemotactic or					
CC	chemokine/chemokine activity, haemostatic or thrombolytic activity,					
CC	receptor/ligand activity, anti-inflammatory activity, tumour					
CC	inhibition activity, or other activities. No evidence of any of					
CC	these activities is given in the specification.					
XX						
SQ	Sequence 1778 BP: 514 A; 386 C; 422 G; 456 T; 0 other;					
QY	Query Match	99.8%	Score 1340.8	DB 18	Length 1778	
Db	Best Local Similarity	99.9%	Pred. No. of			
	Matches 1342	Conservative 0	Mismatches 2	Indels 0	Gaps 0	
QY	1 AAGAATGGCATCTCTTAAGAGAGCTTTTGAAGAAATTAAGAAAGAAATAGCAGCTGTGGA	60				
Db	89 AAGATGGCATCTCTTAAGAGAGCTTTTGAAGAAATTAAGAAAGAAATAGCAGCTGTGGA	148				
QY	61 GATGTGGCTAAAGCAATCATCAACCTGACCTGTTATGGTAAAGCCAGAACAGATCTTAT	120				
Db	149 GATGTGGCTAAAGCAATCATCAACCTGACCTGTTATGGTAAAGCCAGAACAGATCTTAT	208				
QY	121 GAGGAGTTGGACCTTCTGTGTGATACACTGTTGGAGCCAGACTGAGTGGTCCAGAACTTA	180				
Db	209 GAGGAGTTGGACCTTCTGTGTGATACACTGTTGGAGCCAGACTGAGTGGTCCAGAACTTA	268				
QY	181 GAAAAAGCCATCCAAATTAATATGATACCAAAACCTGAGCAGAAAGATGGGCTGGAAGAAAGTTAC	240				
Db	269 GAAAAAGCCATCCAAATTAATATGATACCAAAACCTGAGCAGAAAGATGGGCTGGAAGAAAGTTAC	328				
QY	241 CTGGAGCCAGTGAAGATATCCCACTGGAGAGAGGAGAGAAAGATCACTGTGATGCTGGAG	300				
Db	329 CTGGAGCCAGTGAAGATATCCCACTGGAGAGAGGAGAGAAAGATCACTGTGATGCTGGAG	388				
QY	301 CCAAGAAATTCATAAGATAGCCATCTGTGGCTTTGGCAGCAGACATTTGGACCTTCACAGAA	360				
Db	389 CCAAGAAATTCATAAGATAGCCATCTGTGGCTTTGGCAGCAGACATTTGGACCTTCACAGAA	448				
QY	361 GGCAATTCAGAGAGAAGTTCTGGTGGTGAACCTCTTTCATGTAACCTGAGAGAAAGGCGCTCA	420				
Db	449 GGCAATTCAGAGAGAAGTTCTGGTGGTGAACCTCTTTCATGTAACCTGAGAGAAAGGCGCTCA	508				
QY	421 GAAACCAAGAGGGAAGATGTTGTTTATTAACCAACCTTACATCACTCAAGACGCGTG	480				
Db	509 GAAACCAAGAGGGAAGATGTTGTTTATTAACCAACCTTACATCACTCAAGACGCGTG	568				
QY	481 CAATACCGAAGCAGGAGGAGCGGTGAAAGCTGCCAAGGTGGGAGCTTTGGACTTCTCAT	540				
Db	569 CAATACCGAAGCAGGAGGAGCGGTGAAAGCTGCCAAGGTGGGAGCTTTGGACTTCTCAT	628				
QY	541 CGATCCGAGGCGCTCTCTCCATCTTAAGTCCCTCAACAGGTATTTCAAGAAATACAGGAT	600				
Db	629 CGATCCGAGGCGCTCTCTCTCCATCTTAAGTCCCTCAACAGGTATTTCAAGAAATACAGGAT	688				
QY	601 GGCGGTGCCAAATATTCACACAGCCTGTATTAACGCTGGAAAGATGACAAATATGATGTCAAGA	660				
Db	689 GGCGGTGCCAAATATTCACACAGCCTGTATTAACGCTGGAAAGATGACAAATATGATGTCAAGA	748				
QY	661 ATGACCTTCTCATGGGATCAAAATTTGTCACTTCAGCTAAAGATGGGAGCAGAAAGCTTACCCA	720				
Db	749 ATGACCTTCTCATGGGATCAAAATTTGTCACTTCAGCTAAAGATGGGAGCAGAAAGCTTACCCA	808				
QY	721 GATACCTGATTCCTTCAACACTGTAGCAGAGATCACTGGAGACAAATATCCAGAACAGGTT	780				
Db	809 GATACCTGATTCCTTCAACACTGTAGCAGAGATCACTGGAGACAAATATCCAGAACAGGTT	868				
QY	781 GTACTGTGACATGTGACATCTGACACAGCTGGAGATGTTGGGACAGGGTGCCATGATGATGAC	840				
Db	869 GTACTGTGACATGTGACATCTGACACAGCTGGAGATGTTGGGACAGGGTGCCATGATGATGAC	928				
QY	841 GGTGAGAGCTTTTATATCATGGGAAGCACTCTCACTTATTTAAAGATCTTGGGCTCGTCA	900				

Db 925 GGTGGAGCTTTATATCATGGAGCACTCTCATTTATTAAGATCTTGCGGTGCTCA 988  
 Qy 901 AAGAGACTCTGCGGCTGCTGCTGCACTGCAAGAAACAGTGGGTTGTCCTTC 960  
 Db 989 AAGGAGCTCTGCGGCTGCTGCTGCACTGCAAGAAACAGTGGGTTGTCCTTC 1048  
 Qy 961 CAGTATTATCATGTTACACAGGTAATATTTCCACTACAGTCTGATGAGTCTGAC 1020  
 Db 1049 CAGTATTATCATGTTACACAGGTAATATTTCCACTACAGTCTGATGAGTCTGAC 1108  
 Qy 1021 GCAGGAACCTTTTATCCCACTGCGGCTGCAATTCAGTGGAGTGAAGGCCAGCCATC 1080  
 Db 1109 GCAGGAACCTTTTATCCCACTGCGGCTGCAATTCAGTGGAGTGAAGGCCAGCCATC 1168  
 Qy 1081 ATGAGAGGTTATGAGGCTGCTGCAAGCCCTCAATATCATCTAGTCTTCAAGCATGGA 1140  
 Db 1169 ATGAGAGGTTATGAGGCTGCTGCAAGCCCTCAATATCATCTAGTCTTCAAGCATGGA 1228  
 Qy 1141 GAAGGAGCAGACATCACTTTTGTGATCCAGCTGAGTGGCCAGTCTACTTAT 1200  
 Db 1229 GAAGGAGCAGACATCACTTTTGTGATCCAGCTGAGTGGCCAGTCTACTTAT 1288  
 Qy 1201 GACTTATACAGATATTTCTTCCATCAGTCCACGAGACACCATGACTGTATGAT 1260  
 Db 1289 GACTTATACAGATATTTCTTCCATCAGTCCACGAGACACCATGACTGTATGAT 1348  
 Qy 1261 CCAAGAGCATGATATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC 1320  
 Db 1349 CCAAGAGCATGATATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC 1408  
 Qy 1321 ATGAGAAAGATGCTGCTAGGTCC 1344  
 Db 1409 ATGAGAAAGATGCTGCTAGGTCC 1432

## RESULT 5

AA298034  
 ID AA298034 standard; cDNA; 1863 BP.

AA298034;

09-MAY-2000 (first entry)

Human secreted protein encoding nucleotide sequence SRQ ID NO:28.

Human; secreted protein; diagnosis; cytostatic; immunosuppressive;  
 antiinflammatory; nootropic; neuroprotective; antiallergic; cancer;  
 tumour; neurodegenerative disorder; developmental abnormality; allergy;  
 foetal deficiency; blood disorder; immune system disorder; arthritis;  
 autoimmune disease; hepatic disease; renal disease; inflammation;  
 Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;  
 infection; AIDS; spinal cord injury; transplant rejection; diabetes;  
 asthma; sepsis; acne; psoriasis; cardiovascular disorder;  
 reproductive disorder; gastrointestinal disorder; respiratory disorder;  
 metabolic disorder; food additive; preservative; ss.

Homio sapiens.

WO200004140-A1.

27-JAN-2000.

14-JUL-1999; 99WO-US15849.

15-JUL-1998; 98US-0092921.

15-JUL-1998; 98US-0092922.

15-JUL-1998; 98US-0092956.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y,  
 Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE;

PI Mucenaki M, Endress GA, Soppet DR;  
 XX WPI: 2000-161128/14.  
 DR P-PSDB: AA187081.  
 PT New isolated human genes, useful for diagnosis and treatment of, e.g.  
 cancers, neurological or blood disorders  
 XX Claim 1; Page 319; 494pp; English.

The polynucleotide sequences given in AA298017 to AA298108 encode the  
 human secreted proteins given in AA187064 to AA187223. Human secreted  
 protein can have activities based on the tissues and cells the genes are  
 expressed in. Examples of activities include: cytostatic;  
 immunosuppressive; antiinflammatory; nootropic; neuroprotective; and  
 antiallergic. The polynucleotides and their corresponding secreted  
 polypeptides are useful for preventing, treating or ameliorating medical  
 conditions, e.g. by protein or gene therapy. Also pathological conditions  
 can be diagnosed by determining the amount of the new polypeptides in a  
 sample or by determining the presence of mutations in the new  
 polynucleotides. Human secreted proteins and their polynucleotides can  
 be used for developing products for the diagnosis or treatment of cancer,  
 tumours, neurodegenerative disorders, developmental abnormalities and  
 foetal deficiencies, blood disorders, diseases of the immune system,  
 autoimmune diseases, hepatic and renal disease, inflammation,  
 allergies, Alzheimer's disease, behavioural disorders, schizophrenia,  
 osteoporosis, arthritis, infections, AIDS, spinal cord injuries,  
 transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,  
 cardiovascular disorders, reproductive disorders, gastrointestinal  
 disorders, respiratory disorders and metabolic disorders. The  
 proteins or polynucleotides can also be used as food additives or  
 preservatives. The proteins are also useful for identifying their  
 binding partners. AA298008 to AA298016 and AA187063 are sequence used in  
 the exemplification of the present invention.

Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Query Match 99.8%; Score 1340.8; DB 21; Length 1863;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1342; Conservative 0; Gaps 0;

Qy 1 AAGATGCACTCTTAAGAGACTTTGAAGAAATAAAGAAATAGCCAGCTGTGA 60  
 Db 171 AAGATGCACTCTTAAGAGACTTTGAAGAAATAAAGAAATAGCCAGCTGTGA 230  
 Qy 61 GATGTTGCTAAGCAATCATCACTAGCTGTTATAGTAAAGCCAGAAAGATCTAT 120  
 Db 231 GATGTTGCTAAGCAATCATCACTAGCTGTTATAGTAAAGCCAGAAAGATCTAT 290  
 Qy 121 GAGGATGGAACCTTGTGTTGATGCTGTTGGAACCCAGACTGATGCTCCAGAACCTA 180  
 Db 291 GAGGATGGAACCTTGTGTTGATGCTGTTGGAACCCAGACTGATGCTCCAGAACCTA 350  
 Qy 181 GAAAGAGCATCCAAATATATATACCAAAACCTGAGCAAGATGGCTGAGAAAGTTTAC 240  
 Db 351 GAAAGAGCATCCAAATATATATACCAAAACCTGAGCAAGATGGCTGAGAAAGTTTAC 410  
 Qy 241 CTGAGCCAGTGAATAATACCCACTGAGAGAGGGGAAAGAAATAGCTGATCTGAG 300  
 Db 411 CTGAGCCAGTGAATAATACCCACTGAGAGAGGGGAAAGAAATAGCTGATCTGAG 470  
 Qy 301 CCAAGAAATTCATAGATAGCCATCTGGGCTTGGAGAGCATTTGGGACTCTCCAGAA 360  
 Db 471 CCAAGAAATTCATAGATAGCCATCTGGGCTTGGAGAGCATTTGGGACTCTCCAGAA 530  
 Qy 361 GGCATTACAGCAAGTCTGTTGATGACCTTTTCATGATGACAGAGAGGCGCTCA 420  
 Db 531 GGCATTACAGCAAGTCTGTTGATGACCTTTTCATGATGACAGAGAGGCGCTCA 590  
 Qy 421 GAAGCAAGAGGAAAGTTGTTGTTATTAACCACTTACATCAACTCTCAAGAGCGTG 480  
 Db 591 GAAGCAAGAGGAAAGTTGTTGTTATTAACCACTTACATCAACTCTCAAGAGCGTG 650

QY 481 CAATACGACGACGAGGCGGCTGGAGAGCTCCAAAGTGGGGCTTTGGCATCTCTCAT 540  
 Db 651 CAATACGACGACGAGGCGGCTGGAGAGCTCCAAAGTGGGGCTTTGGCATCTCTCAT 710  
 QY 541 CGATCCGCGGCT 600  
 Db 711 CGATCCGCGGCT 770  
 QY 601 GCGCGTCCCAAAATTCACAGAGCTGTATTAAGGTTGAGAAATGCAAAATGATGCAAGA 660  
 Db 771 GCGCGTCCCAAAATTCACAGAGCTGTATTAAGGTTGAGAAATGATGCAAGA 830  
 QY 661 ATGGCTTCTCATGGGATCAAAATTTGTCAATGAGCTAAAGATGGGGCAAGACCTACCCA 720  
 Db 831 ATGGCTTCTCATGGGATCAAAATTTGTCAATGAGCTAAAGATGGGGCAAGACCTACCCA 890  
 QY 721 GATCTGATTTCTTCAACACTGTGTAGACAGATCACTGGGAGCAAAATATCCAGAACGTT 780  
 Db 891 GATCTGATTTCTTCAACACTGTGTAGACAGATCACTGGGAGCAAAATATCCAGAACGTT 950  
 QY 781 GTACTGGTCAGTGGACATCTGGAAGAGCTGGGATGTTGGGAGGCGCATGGATGATGAC 840  
 Db 951 GTACTGGTCAGTGGACATCTGGAAGAGCTGGGATGTTGGGAGGCGCATGGATGATGAC 1010  
 QY 841 GGTGAGGCTTTATATCATGAGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCA 900  
 Db 1011 GGTGAGGCTTTATATCATGAGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCA 1070  
 QY 901 AAGAGGACTCTGGGCTGTGTCTCTGAGCTGACAGAGAAACAAGGTGAGTGTGCTTC 960  
 Db 1071 AAGAGGACTCTGGGCTGTGTCTCTGAGCTGACAGAGAAACAAGGTGAGTGTGCTTC 1130  
 QY 961 CAGTATTATCAGTTACACAAAGTAAATTTTCCAACTACAGTCTGTGTGATGAGAGCTGAC 1020  
 Db 1131 CAGTATTATCAGTTACACAAAGTAAATTTTCCAACTACAGTCTGTGTGATGAGAGCTGAC 1190  
 QY 1021 GCAGGAACCTTTCTTACCACCTGAGGCTGCATTCATCTGAGAGTAAAGGCCAGGCCATC 1080  
 Db 1191 GCAGGAACCTTTCTTACCACCTGAGGCTGCATTCATCTGAGAGTAAAGGCCAGGCCATC 1250  
 QY 1081 ATGAGAGGTTTATGAGGCTGTGTCTGAGAGCTCAATATCACTCAGTCTTGAAGCCATGA 1140  
 Db 1251 ATGAGAGGTTTATGAGGCTGTGTCTGAGAGCTCAATATCACTCAGTCTTGAAGCCATGA 1310  
 QY 1141 GAAGGAGCAGACATCACTTTGATCCAGCTGAGAGCTGAGAGCTCACTTATGAT 1200  
 Db 1311 GAAGGAGCAGACATCACTTTGATCCAGCTGAGAGCTGAGAGCTCACTTATGAT 1370  
 QY 1201 GACTTATACAGTATTTCTTCCATCACTCCACGAGACACCATGATGATGAT 1260  
 Db 1371 GACTTATACAGTATTTCTTCCATCACTCCACGAGACACCATGATGATGAT 1430  
 QY 1261 CCAAGCAGATGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC 1330  
 Db 1431 CCAAGCAGATGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC 1490  
 QY 1321 ATGAGAGAAATGCTGCTGATGCTCC 1344  
 Db 1491 ATGAGAGAAATGCTGCTGATGCTCC 1514

## RESULT 6

AAD11647 standard; cDNA; 1863 BP.

AAD11647;

24-SEP-2001 (first entry)

Human secreted protein-encoding gene 18 cDNA clone HRA0315, SEQ ID NO:28.  
 Human; secreted protein; proliferative disorder; cancer; tumour; asthma;  
 foetal abnormality; developmental abnormality; haematopoietic disorder;

KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;  
 KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;  
 KW inflammation; neurological disorder; Alzheimer's disease; food additive;  
 KW angiotensin disorder; kidney disorder; gastrointestinal disorder; allergy;  
 KW pregnancy-related disorder; endocrine disorder; infection; wound healing;  
 KW cell culture; chemotaxis; vulnery; binding partner identification;  
 KW gene therapy; ss.  
 KW OS  
 KW Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 99..1517  
 FT /\*tag= a  
 FT /product= "Human secreted protein precursor"  
 FT sig\_peptide 99..170  
 FT /\*tag= b  
 FT mat\_peptide 171..1514  
 FT /\*tag= c  
 FT /product= "Mature human secreted protein"  
 XX  
 EN W0200151504-A1.  
 XX  
 PD 19-JUL-2001.  
 XX  
 PF 12-JAN-2001; 2001WO-US00911.  
 XX  
 PR 13-JAN-2000; 2000US-0482273.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Komatsu JI GA, Duan DR, Rosen CA, Moore PA, Shi Y;  
 PI Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;  
 PI Endress GA, Muscenski M, Ebner R;  
 XX  
 DR WP1; 2001-425865/45.  
 XX  
 PT Isolated nucleic acid molecule encoding a human secreted protein is  
 PT used in preventing, treating or ameliorating a medical condition -  
 XX  
 Claim 1; Page 686; 864pp; English.

AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted  
 CC protein genes, and AAB06041-AAB06132 represent the proteins they encode.  
 CC AAB06133-AAB06205 represent human secreted protein fragments.  
 CC The secreted proteins and their genes are useful for preventing, treating  
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
 CC Pathological conditions can be diagnosed by determining the amount of the  
 CC new protein in a sample or by determining the presence of mutations in  
 CC the new genes. Specific uses are described for each of the 71 genes,  
 CC based on the tissues in which they are most highly expressed, and include  
 CC developing products for the diagnosis or treatment of proliferative  
 CC disorders, cancer, tumours, foetal and developmental abnormalities,  
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,  
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,  
 CC pregnancy-related disorders, endocrine disorders, and infections. The  
 CC proteins can also be used to aid wound healing and epithelial cell  
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
 CC before transplantation, for supporting cell culture of primary tissues,  
 CC to regenerate tissues, to identify their cognate ligands or binding  
 CC partners, and in chemotaxis, and can be used as a food additive or  
 CC preservative to modify storage properties. Antibodies specific for a  
 CC protein of the invention can be used in alleviating symptoms associated  
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).  
 CC The present sequence represents a human secreted protein-encoding cDNA of  
 CC the invention.



**SQ** Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Query Match	99.8%;	Score 1340.8;	DB 22;	Length 1863;
-------------	--------	---------------	--------	--------------

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY	1	AAGATGGCATCTCTTAAGAGGACTTTTGAAAGAAATTAAGAAAGAAATTAAGCATCTGTGGA	60
Db	171	AAGATGGCATCTCTTAAGAGGACTTTTGAAAGAAATTAAGAAAGAAATTAAGCATCTGTGGA	230
OY	61	GATGTGGTAAAGCAATCATCAACCTAGCTGTTATATGGTAAAGCCACAGACAGATCTAT	120
Db	231	GATGTGGTAAAGCAATCATCAACCTAGCTGTTATATGGTAAAGCCACAGACAGATCTAT	290
OY	121	GAGCATTTGGCACTTCTGGTTGATATCTGTTGGACCCGAGCTGAGTGGCTCCAGAACCTTA	180
Db	291	GAGCATTTGGCACTTCTGGTTGATATCTGTTGGACCCGAGCTGAGTGGCTCCAGAACCTTA	350
OY	181	GAAAAAGGCATCCAAATTAATGTAACAAAACCTGGACGAAAGATGGGCTGGAGAAATTTAC	240
Db	351	GAAAAAGGCATCCAAATTAATGTAACAAAACCTGGACGAAAGATGGGCTGGAGAAATTTAC	410
OY	241	CTGGAGCCAGTGGAGATACCCCACTGGAGAGAGGAGAAAGATTCAGCTGTGATGTGTGAG	300
Db	411	CTGGAGCCAGTGGAGATACCCCACTGGAGAGAGGAGAAAGATTCAGCTGTGATGTGTGAG	470
OY	301	CCAGAAATTCATTAAGATAGCCATCTGGGTCTTGGCAGCAGATTTGGGACTCTTCCAGAA	360
Db	471	CCAGAAATTCATTAAGATAGCCATCTGGGTCTTGGCAGCAGATTTGGGACTCTTCCAGAA	530
OY	361	GGCATTTACAGCAGAGATCTGGTGGTGACCTCTTTGATGAATCTCAGAGAGGGCTCTCA	420
Db	531	GGCATTTACAGCAGAGATCTGGTGGTGACCTCTTTGATGAATCTCAGAGAGGGCTCTCA	590
OY	421	GAAGCAAGAGGAGAGATTTGTTTATTAACAACCTTACATCACTAACAGACGGTG	480
Db	591	GAAGCAAGAGGAGAGATTTGTTTATTAACAACCTTACATCACTAACAGAGCGGTG	650
OY	481	CAATCCGAAACGACAGGGGGCGGTGGAAAGTGCAGAGTGGGGGCTTGGCATCTCTCAAT	540
Db	651	CAATCCGAAACGACAGGGGGCGGTGGAAAGTGCAGAGTGGGGGCTTGGCATCTCTCAAT	710
OY	541	CGATCCGTGGGCTCTCTTCTCCATCTACAGTCTCTCAACACAGTATTCAGGAATACCAAGAT	600
Db	711	CGATCCGTGGGCTCTCTTCTCCATCTACAGTCTCTCAACACAGTATTCAGGAATACCAAGAT	770
OY	601	GGCGTGCCCAAAATTCCAACAGCTGTATTAACGTGGAAAGATGCAGAAATGATGTCAAGA	660
Db	771	GGCGTGCCCAAAATTCCAACAGCTGTATTAACGTGGAAAGATGCAGAAATGATGTCAAGA	830
OY	661	ATGGCTTCTCANTGGGATCAAAATTTGTCATTCAGCTTAAGATGGGGGCAAAAGCTTAACCA	720
Db	831	ATGGCTTCTCANTGGGATCAAAATTTGTCATTCAGCTTAAGATGGGGGCAAAAGCTTAACCA	890
OY	721	GATACGTATTCCTTCAACACTGTACAGAGATCACTGGGAGCAAAATATTCAGAACAGATT	780
Db	891	GATACGTATTCCTTCAACACTGTACAGAGATCACTGGGAGCAAAATATTCAGAACAGATT	950
OY	781	GTACTGTGTCAGTGGACATCTGGACACGTGGGATGTTGGGCAAGGGTGCCATGATGATGGC	840
Db	951	GTACTGTGTCAGTGGACATCTGGACACGTGGGATGTTGGGCAAGGGTGCCATGATGATGGC	1010
OY	841	GGTGAAGCCTTTATATTCATGGGAGACACTCACTTAATTAAGATCTTGGGCTGGGCTCCA	900
Db	1011	GGTGAAGCCTTTATATTCATGGGAGACACTCACTTAATTAAGATCTTGGGCTGGGCTCCA	1070
OY	901	AAAGAGACTCTGCGGCTGGTGTCTGTGACTGCAAGAAAGCAAGGTGAGATTGTCCTTTT	960
Db	1071	AAAGAGACTCTGCGGCTGGTGTCTGTGACTGCAAGAAAGCAAGGTGAGATTGTCCTTTT	1130
OY	961	CAGTATTTATCACTTACACAAAGTAAATATTTTCCAACTACAGTCTGTGATGGAAGTCTGAC	1020
Db	1131	CAGTATTTATCACTTACACAAAGTAAATATTTTCCAACTACAGTCTGTGATGGAAGTCTGAC	1190

QY	1021	GCAGAACTCTTACCACTGGCTGCAATTCACTGGCATGAAAGGCCAGGCCATC	1080
Db	1191	GCAGAACTCTTACCACTGGCTGCAATTCACTGGCATGAAAGGCCAGGCCATC	1250
QY	1081	ATGAGGAGGTATTAGAGCTGCTGCAGCCCTCAATATCACTGAGTCTGAGCATGGA	1140
Db	1261	ATGAGGAGGTATTAGAGCTGCTGCAGCCCTCAATATCACTGAGTCTGAGCATGGA	1310
QY	1141	GAAGGACAGACATCAACTTTTGGATCCAACTGAGTGCCTGAGGCAGTCTACTTGAT	1200
Db	1311	GAAGGACAGACATCAACTTTTGGATCCAACTGAGTGCCTGAGGCAGTCTACTTGAT	1370
QY	1201	GACTTATTCAGATATTTCTTCTTCATCACTCCACGAGACACCATGACTGTCATGAT	1260
Db	1371	GACTTATTCAGATATTTCTTCTTCATCACTCCACGAGACACCATGACTGTCATGAT	1430
QY	1261	CCAAAGCAGATGATATGTTGCTGCTGCTGTTGGGCTGTGTTTCTTATGTTGTCAGAC	1320
Db	1431	CCAAAGCAGATGATATGTTGCTGCTGCTGTTGGGCTGTGTTTCTTATGTTGTCAGAC	1490
QY	1321	ATGGAAGAAATGCTGCCAGGTCC	1344
Db	1491	ATGGAAGAAATGCTGCCAGGTCC	1514

## RESULT 7

ID ABK69743 standard; cDNA; 1863 BP.

AC ABK69743;

DT , 15-JUL-2002 (first entry)

DE Human secreted protein gene 18 #1

Human, B; gene, secreted protein; gene therapy; immunosuppressive;  
antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiac;  
vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;  
viral; fungicide; ophthalmologic; autoimmune disease; neoplasm;  
rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;  
cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;  
angiogenesis; nervous system disorders; Alzheimer's disease; infection;  
ocular disorder; corneal infection; wound healing; skin aging;  
epithelial cell proliferation; food additive.

OS Homo sapiens.

PN W0200226931-A2.

PD 04-APR-2002

PF 24-SEP-2001; 2001WO-US29871.

PR 25-SEP-2000; 2000US-234925P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y;  
PI Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR  
PI Endress GA, Mucenski M, Ebner R;

DR WPI; 2002-362489/39

DR P-PSDB; ABG33880

PT Novel 71 isolated secreted polypeptides and polynucleotides encoding  
PT the polypeptides, useful for treating Huntington's disease, sepsis,  
PT meningitis, thrombocytopaenia, haemolytic anaemia, rheumatoid arthritis  
PT asthma -

Claim 1; Page 1170; 1478pp; English.



CC The invention relates to an isolated nucleic acid molecule (or its  
 CC fragment, homologue complement or allele variant) encoding a human  
 CC secreted protein (and its fragment, domain, epitope, variant, secreted  
 CC form and species variant). Also included are a recombinant vector  
 CC comprising the nucleic acid, a recombinant host cell comprising the  
 CC vector, an antibody against the secreted protein, a recombinant host cell  
 CC that expresses the secreted protein and a method of identifying a binding  
 CC partner of the secreted protein. The nucleic acid and protein are used to  
 CC prevent, diagnose, treat or ameliorate a medical condition in e.g.  
 CC human, mice, rabbits, goats, horses, cats, dogs, chickens or sheep  
 CC for example autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischaemia, angiodenesis, nervous system disorders e.g.  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
 CC ocular disorders e.g. corneal infection. Many other diseases and  
 CC disorders are listed in the specification. The polypeptides can also be  
 CC used to aid wound healing an epithelial cell proliferation, to prevent  
 CC skin aging due to sunburn, to maintain organs before transplantation, for  
 CC supporting cell culture of primary tissues, to regenerate tissues and in  
 CC chemotaxis. The polypeptides can also be used as a food additive or  
 CC preservative to increase or decrease storage capabilities. The present  
 CC sequence encodes a novel human secreted protein of the invention.

XX Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Query Match 99.8%; Score 1340.8; DB 24; Length 1863;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;

Matches 134; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAGATGGCATCTCTAAGAGACCTTTGAGAAATTAAGAAATAGCCAGCTGTGA 60  
 DB 171 AAGATGGCATCTCTAAGAGACCTTTGAGAAATTAAGAAATAGCCAGCTGTGA 230  
 QY 61 GATTTGCTAAGCAATCATATACCTAGCTTTATAGTAAAGCCAGACATCTAT 120  
 DB 231 GATTTGCTAAGCAATCATATACCTAGCTTTATAGTAAAGCCAGACATCTAT 290  
 QY 121 GAGGATTTGGCATCTCTGTTGATAGTGTGACCCAGACTGAGTGGCTCCAAACCTA 180  
 DB 291 GAGGATTTGGCATCTCTGTTGATAGTGTGACCCAGACTGAGTGGCTCCAAACCTA 350  
 QY 181 GAAAAAGCATCCAAATTTATGTACAAAACCTGAGCAAGATGGCTGAGAAAGTTTAC 240  
 DB 351 GAAAAAGCATCCAAATTTATGTATACAAAACCTGAGCAAGATGGCTGAGAAAGTTTAC 410  
 QY 241 CTGAGCCAGTGAATACCCCACTGAGAGAGGAGAAAGATCACTGTGATCTGGAG 300  
 DB 411 CTGAGCCAGTGAATACCCCACTGAGAGAGGAGAAAGATCACTGTGATCTGGAG 470  
 QY 301 CCAAGATTTATAGATAGCATCTGGGTGTTGGACAGACATTTGGGATCTCTCCAGAA 360  
 DB 471 CCAAGATTTATAGATAGCATCTGGGTGTTGGACAGACATTTGGGATCTCTCCAGAA 530  
 QY 361 GGCATTACAGAGAGATTTCTGTGTGACCTTTTCATGTAATGACAGAAAGGCTTCA 420  
 DB 531 GGCATTACAGAGAGATTTCTGTGTGACCTTTTCATGTAATGACAGAAAGGCTTCA 590  
 QY 421 GAAGCAAGAGAGATTTGTTTATTAACCACTTATCATCACTCAAGAGCGGTG 480  
 DB 591 GAAGCAAGAGAGATTTGTTTATTAACCACTTATCATCACTCAAGAGCGGTG 650  
 QY 481 CAATACCGAAGCGAGGGGGCGGTGGAAGCTGCCAAGTGGGGGCTTTGGCATCTCTCAT 540  
 DB 651 CAATACCGAAGCGAGGGGGCGGTGGAAGCTGCCAAGTGGGGGCTTTGGCATCTCTCAT 710  
 QY 541 CGATCCGTGGCTCTCTTCATCTACAGTCTCTCAACAGGATTTCAAGAAATACAGAT 600  
 DB 711 CGATCCGTGGCTCTCTTCATCTACAGTCTCTCAACAGGATTTCAAGAAATACAGAT 770  
 QY 601 GGCCTGCCAAATTTCCAAAGCTGTATTAAGGTGAAGATGCAAAATGATGTCAAGA 660  
 DB 771 GGCCTGCCAAATTTCCAAAGCTGTATTAAGGTGAAGATGCAAAATGATGTCAAGA 830

QY 661 ATGGCTTCTCATGGGATCAAAATTTGTCATTAGCTAAAGATGGGGGCAAGACTTACCA 720  
 DB 831 ATGGCTTCTCATGGGATCAAAATTTGTCATTAGCTAAAGATGGGGGCAAGACTTACCA 890  
 QY 721 GATTCATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAACAGTT 780  
 DB 891 GATTCATGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAACAGTT 950  
 QY 781 GTACTGTCAGTGAACATCTGACACAGTGGGATCTTGGGAGAGGTGCATGATGATGCG 840  
 DB 951 GTACTGTCAGTGAACATCTGACACAGTGGGATCTTGGGAGAGGTGCATGATGATGCG 1010  
 QY 841 GGTGAGACCTTTTATATCATGGGAGAGCACTCTCACTTATTAAGATCTTGGGCTCGTCCA 900  
 DB 1011 GGTGAGACCTTTTATATCATGGGAGAGCACTCTCACTTATTAAGATCTTGGGCTCGTCCA 1070  
 QY 901 AAGAGACTGTCGGCTGTGTCTGTGACCTGCAGAGAACAGAGTGGAGTTGCTTC 960  
 DB 1071 AAGAGACTGTCGGCTGTGTCTGTGACCTGCAGAGAACAGAGTGGAGTTGCTTC 1130  
 QY 961 CAGTATTTACATTAACCAAGATTAATTTTCCATCACTGATGATGATGATGATGATGAT 1020  
 DB 1131 CAGTATTTACATTAACCAAGATTAATTTTCCATCACTGATGATGATGATGATGATGAT 1190  
 QY 1021 GCAAGAACCTTTTACCACTGAGGCTGCAATTCATGAGTGAAGAGGAGGAGGAGGAGGAG 1080  
 DB 1191 GCAAGAACCTTTTACCACTGAGGCTGCAATTCATGAGTGAAGAGGAGGAGGAGGAGGAG 1250  
 QY 1081 ATGAGAGAGGTTATGAGCTGCTGACAGCCCTCAATATCACTGAGTCTGAGGATGGA 1140  
 DB 1251 ATGAGAGAGGTTATGAGCTGCTGACAGCCCTCAATATCACTGAGTCTGAGGATGGA 1310  
 QY 1141 GAAGGACACACATCAACTTTTGGATCAAGCTGAGTGGCTTGGAGCCAGTCTAATTGAT 1200  
 DB 1311 GAAGGACACACATCAACTTTTGGATCAAGCTGAGTGGCTTGGAGCCAGTCTAATTGAT 1370  
 QY 1201 GACTTATACAGATATTTCTTCCATCACTCCCAAGAGAGACCAATGATGATGATGAT 1260  
 DB 1371 GACTTATACAGATATTTCTTCCATCACTCCCAAGAGAGACCAATGATGATGATGAT 1430  
 QY 1261 CCAAGCAGATGATGTTGCTGCTGTGTGTTGGGCTGTGTTCTTATGTTGTCAGAC 1320  
 DB 1431 CCAAGCAGATGATGTTGCTGCTGTGTGTTGGGCTGTGTTCTTATGTTGTCAGAC 1490  
 QY 1321 ATGAGAGAAATGCTGCTAGGCTC 1344  
 DB 1491 ATGAGAGAAATGCTGCTAGGCTC 1514

RESULT 8  
 ACCS0817  
 ID ACCS0817 standard; cDNA; 1863 BP.  
 XX ACCS0817;  
 AC ACCS0817;  
 DT 12-JUN-2003 (first entry)  
 XX  
 DE Human secreted protein coding sequence, SEQ ID 484.  
 XX  
 KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cybostatic;  
 KW valnetary; antiinflammatory; nootropic; neuroprotective;  
 KW antiparkinsonian; gene therapy; human; cardiovascular disorder;  
 KW gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200295010-A2.  
 XX  
 PD 28-NOV-2002.  
 XX  
 PF 19-MAR-2002; 2002WO-US09785.  
 XX



DT 04-APR-2003 (first entry)  
XX Secreted protein-encoding gene 142 cDNA clone HRACJ35, SEQ ID NO:274.  
XX  
XX Human; secreted protein; digestive disorder; gastrointestinal disorder;  
XX mouth; oesophagus; stomach; small intestine; large intestine; liver;  
XX biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;  
XX immune disorder; inflammation; infection; wound healing; drug screening;  
XX chromosome identification; chromosome mapping; cytostatic; gene therapy;  
XX antiinflammatory; immunosuppressive; vulnerary; chromosome 8q22.2;  
XX gene; ss.  
XX Homo sapiens.  
XX WO200276488-A1.  
XX  
XX 03-OCT-2002.  
XX  
XX 19-MAR-2002; 2002WO-US08276.  
XX  
XX 21-MAR-2001; 2001US-277340P.  
XX 19-JUL-2001; 2001US-306171P.  
XX 13-NOV-2001; 2001US-331287P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
XX WPI, 2003-029900/02.  
XX P-PSDB; ABR00274.  
XX  
XX New human secreted proteins and nucleic acids, useful for detecting,  
XX preventing, diagnosing, prognosticating, treating and/or ameliorating,  
XX e.g. gastrointestinal diseases and disorders, or cancers -  
XX  
XX Claim 21; Page 918; 1216pp; English.  
XX  
XX AB271190-AB271478 represent cDNAs corresponding to 178 human secreted  
XX protein genes, and ABP00011-ABP00299 represent the proteins they encode.  
XX AB271479-AB271540 represent human secreted protein genomic fragments. The  
XX invention also encompasses antibodies specific for the secreted proteins,  
XX the use of the secreted proteins in drug screening, and recombinant  
XX vectors and host cells comprising a nucleic acid of the invention. The  
XX secreted proteins, nucleic acids encoding them, antibodies or antibody  
XX fragments specific for the secreted proteins, and modulators of protein  
XX activity are useful for diagnosing, treating, ameliorating or preventing  
XX digestive disorders. Such conditions include disorders of the mouth,  
XX oesophagus, stomach, small intestine, large intestine, liver, biliary  
XX tract and pancreas, and include cancers of these organs and tissues. The  
XX secreted proteins and their nucleic acids may also be used in the  
XX treatment of immune disorders, inflammation, infection,  
XX hyperproliferative disorders, and to promote wound healing. Nucleic acids  
XX of the invention may be used for chromosome identification, chromosome  
XX mapping, in gene therapy, for identifying individuals from minute  
XX biological samples, as hybridisation probes, and as molecular weight  
XX markers. The present sequence represents a human secreted protein-  
XX encoding cDNA clone of the invention.  
XX  
XX Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;  
SQ  
Query Match 99.8%; Score 1340.8; DB 25; Length 1863;  
Best Local Similarity 99.9%; Pred. No. 0; Indels 0; Gaps 0;  
Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AAGAAATGCACTCTTAAGAGGACTTTTGAAGAAATTAAGAAATGAGCCAGCTGTGGA 60  
DB 171 AAGAAATGCACTCTTAAGAGGACTTTTGAAGAAATTAAGAAATGAGCCAGCTGTGGA 230  
QY 61 GATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCCAAGACGATCTTAT 120  
DB 231 GATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCCAAGACGATCTTAT 290  
QY 121 GAGCAATTGCACTTCTGTTGATATCTGTGGAACCCAGACTGATGCTCCAGAACTTA 180

DB 291 GAGCAATTGCACTTCTGTTGATATCTGTGGAACCCAGACTGATGCTCCAGAACTTA 350  
QY 181 GAAAGACCATCCAAATTAATGTAACCAAACTGAGCAAGATGGGCTGAGAAAGTTTAC 240  
DB 351 GAAAGACCATCCAAATTAATGTAACCAAACTGAGCAAGATGGGCTGAGAAAGTTTAC 410  
QY 241 CTGAGGACAGTGAAGATATACCCCTGAGGAGAGGAGAGAAATAGCTGTGATCTGGAG 300  
DB 411 CTGAGGACAGTGAAGATATACCCCTGAGGAGAGGAGAGAAATAGCTGTGATCTGGAG 470  
QY 301 CCAAGAAATTAAGATATGACCATCTGGGCTTTGGACAGCAATGGAGCTCTCCAGAA 360  
DB 471 CCAAGAAATTAAGATATGACCATCTGGGCTTTGGACAGCAATGGAGCTCTCCAGAA 530  
QY 361 GGCATTACAGCAAGATTTGTTGTTGTAACCAACTTACATCACTCAAGAGCGGTG 420  
DB 531 GGCATTACAGCAAGATTTGTTGTTGTAACCAACTTACATCACTCAAGAGCGGTG 590  
QY 421 GAAGCAAGAGGAAATGTTGTTTATTAACCAACTTACATCACTCAAGAGCGGTG 480  
DB 591 GAAGCAAGAGGAAATGTTGTTTATTAACCAACTTACATCACTCAAGAGCGGTG 650  
QY 481 CAATACCGAAACGAGGGGGCGGTGGAAGCTGCCAAGGTGGGGCTTTGGCATCTTCATT 540  
DB 651 CAATACCGAAACGAGGGGGCGGTGGAAGCTGCCAAGGTGGGGCTTTGGCATCTTCATT 710  
QY 541 CGATCCGTCGCTCTCTTCATCTTACATCACTCAAGATTCAGATTAACAGAT 600  
DB 711 CGATCCGTCGCTCTCTTCATCTTACATCACTCAAGATTCAGATTAACAGAT 770  
QY 601 GGGCTGGCCAAATTTCCAGACGCTGATTAACGTTGGAAGTGAAGAAATGATGTCAGA 660  
DB 771 GGGCTGGCCAAATTTCCAGACGCTGATTAACGTTGGAAGTGAAGAAATGATGTCAGA 830  
QY 661 ATGCTCTTCATGAGATCAAAATTTGTCATTGAGCTAAAGATGGGGGCAAGACTTACCA 720  
DB 831 ATGCTCTTCATGAGATCAAAATTTGTCATTGAGCTAAAGATGGGGGCAAGACTTACCA 890  
QY 721 GATATGATTTCTTCAACACTGTGAGAGATCACTGGAGCAAAATTAACAGAGTT 780  
DB 891 GATATGATTTCTTCAACACTGTGAGAGATCACTGGAGCAAAATTAACAGAGTT 950  
QY 781 GTATGTCAGTGAACATCTGAGACGCTGGAGATTTGGGAGGTCATGAGATGGC 840  
DB 951 GTATGTCAGTGAACATCTGAGACGCTGGAGATTTGGGAGGTCATGAGATGGC 1010  
QY 841 GGTGAGCTCTTATATATATGAGGAAAGCACTCTCACTTAATTAAGATCTTGGCTGCTCA 900  
DB 1011 GGTGAGCTCTTATATATATGAGGAAAGCACTCTCACTTAATTAAGATCTTGGCTGCTCA 1070  
QY 901 AAGAGACTCTGCGGCTGTGCTGTGAGCTGACAGAGAAACAGGTGAGTGTGCTTC 960  
DB 1071 AAGAGACTCTGCGGCTGTGCTGTGAGCTGACAGAGAAACAGGTGAGTGTGCTTC 1130  
QY 961 CAGATTAATGATTAACCAAGATTAATTTCCAACTGACCTGTGATGAGATCTGAC 1020  
DB 1131 CAGATTAATGATTAACCAAGATTAATTTCCAACTGACCTGTGATGAGATCTGAC 1190  
QY 1021 GCAGGAACCTTCTTAACCACTGGGCTGCAATTCAGCTGGAGTAAAGCCAGGGCCATC 1080  
DB 1191 GCAGGAACCTTCTTAACCACTGGGCTGCAATTCAGCTGGAGTAAAGCCAGGGCCATC 1250  
QY 1081 ATGAGAGAGGTATGAGCTGTGACAGCCCTCAATATCACTAGTCTTGAGCCATGGA 1140  
DB 1251 ATGAGAGAGGTATGAGCTGTGACAGCCCTCAATATCACTAGTCTTGAGCCATGGA 1310  
QY 1141 GAAAGGACAGACATCACTTTTGAATCCAGCTGGAATGCTGGAAGCAAGCTTCTGAT 1200  
DB 1311 GAAAGGACAGACATCACTTTTGAATCCAGCTGGAATGCTGGAAGCAAGCTTCTGAT 1370  
QY 1201 GACTTAATACAGTATTTCTTCTTCATCACTCCACGAGAGACACATGACTGTATGAT 1260

Db	1371	GACCTATATACAGTATTCTTCTTCATCATCACCACGAGACACATGATCATGGAT	1430
Oy	1261	CCAAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGGACAC	1320
Db	1431	CCAAAGCAGATGAAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGGACAGAC	1490
Oy	1321	ATGGAAGAAATGCTGCTTAGGCC	1344
Db	1491	ATGGAAGAAATGCTGCTTAGGTC	1514
RESULT 10			
ID	AAZ58313	standard; cDNA, 1884 BP.	
AC	AAZ58313;		
XX			
DT	08-MAY-2000	(first entry)	
XX			
DE	Human peptidase NALAD-ase IV cDNA.		
XX			
KM	NALAD-ase IV; N-acetylated alpha-linked acidic dipeptidase; human;		
KM	chromosome 8q21.3; prostate cancer; neurodegenerative disease;		
KM	Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;		
KM	peripheral neuropathy; Huntington's disease; acute brain injury;		
KM	multiple sclerosis; peripheral nerve trauma; ischaemia; dementia;		
KM	gene therapy; diagnosis; nootropic; neuroprotective; neuroleptic;		
XX	antiparkinsonian; anticonvulsant; vasotropic; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	149..1567	
FT		/*tag= a	
XX			
PN	WO200004157-A2.		
XX			
PD	27-JAN-2000.		
XX			
PF	14-JUL-1999;	99WO-GB02241.	
XX			
PR	14-JUL-1998;	98GB-0015284.	
XX			
PA	(JANC ) JANSEN PHARM NV.		
PI	Pangalos M, Neefs JEFM, Peeters DCG;		
XX			
DR	WPI, 2000-182424/16.		
DR	P-PSDB; AAZ58879.		
XX			
PT	New human N-acetylated alpha-linked acidic dipeptidases for treating		
PT	neural disorders e.g. Alzheimer's disease, schizophrenia and		
PT	Parkinson's disease		
XX			
PS	Claim 7, Fig 5, 95pp; English.		
XX			
CC	The present sequence is that of cDNA coding for human		
CC	N-acetylated alpha-linked acidic dipeptidase IV (NALAD-ase IV,		
CC	see AAZ58879). The cDNA was obtained from a gall bladder cDNA		
CC	library. Analysis of the open reading frame predicts a type II		
CC	integral membrane protein with 5 potential N-glycosylation sites.		
CC	The NALAD-ase II gene was mapped to chromosome 8q21.3.		
CC	NALAD-ase IV expression was low in all tissues examined by RT-PCR.		
CC	The invention provides human NALAD-ase I, II and IV polypeptides,		
CC	cDNAs, antisense nucleic acids, vectors, host cells, transgenic		
CC	organisms, antagonists and agonists. These are useful for treating		
CC	neural disorders such as Alzheimer's disease, schizophrenia, ALS,		
CC	Parkinson's disease, peripheral neuropathy, Huntington's disease,		
CC	acute brain injury, multiple sclerosis, exposure to neurotoxins,		
CC	peripheral nerve trauma, ischaemia or dementia (claimed). Nucleic		
CC	acids can also be used for gene therapy and for genetic screening		
CC	of predisposition to disorders associated with NALAD-ase.		
XX			

Seq	Sequence	1884 BP; 520 A; 432 C; 458 G; 474 T; 0 other;
Query Match	99.8%; Score 1340.8; DB 21; Length 1884;	
Best Local Similarity	99.9%; Pred. No. 0;	
Matches 1342; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	
QY	1 AAGAAATGCACTCTTAAGAGCACTTTGAAGAAATTAAGAAGAAATTAAGCAGCTGGA	60
DB	221 AAGAAATGCACTCTTAAGAGCACTTTGAAGAAATTAAGAAGAAATTAAGCAGCTGGA	280
QY	61 GATTTGCTAAAGCAATCATCAACCTAGCTGTTATGTTAAAGCCAGAACAGATCCAT	120
DB	281 GATTTGCTAAAGCAATCATCAACCTAGCTGTTATGTTAAAGCCAGAACAGATCCAT	340
QY	121 GAGCGATTGGCACTTCTGTTGATGATCTGTTGAGCCAGACTGAGTGGCTCCAGAACCTA	180
DB	341 GAGCGATTGGCACTTCTGTTGATGATCTGTTGAGCCAGACTGAGTGGCTCCAGAACCTA	400
QY	181 GAAAAAGCAATCCAAATTAATGTAACCAAACTCGAGCAAGATGGGCTGGAGAAAGTTTAC	240
DB	401 GAAAAAGCAATCCAAATTAATGTAACCAAACTCGAGCAAGATGGGCTGGAGAAAGTTTAC	460
QY	241 CTGGAGCCAGAGAGAAATPACCCCACTGGAGAGGGGAGAGAAATGATGATGCTGGAG	300
DB	461 CTGGAGCCAGAGAGAAATPACCCCACTGGAGAGGGGAGAGAAATGATGATGCTGGAG	520
QY	301 CCAAGAAATTCATTAAGATAGCCATCTGGGCTTTGGAGCAGCATTTGGGACTCTCCAGAA	360
DB	521 CCAAGAAATTCATTAAGATAGCCATCTGGGCTTTGGAGCAGCATTTGGGACTCTCCAGAA	580
QY	361 GGCATTACAGCAAGATTCTGCTGGTGAACCTCTTCCATGAATCTGCAAGAAAGGGCTCTCA	420
DB	581 GGCATTACAGCAAGATTCTGCTGGTGAACCTCTTCCATGAATCTGCAAGAAAGGGCTCTCA	640
QY	421 GAAGCAAGAGGGAAGATTGTTGTTTATTAACCACTTAACATCACTCAAGAACGGTG	480
DB	641 GAAGCAAGAGGGAAGATTGTTGTTTATTAACCACTTAACATCACTCAAGAACGGTG	700
QY	481 CAATACCGCAAGCAGGGGGCGGTGGAACTGCCAAGGTGGGGGCTTTGGACTCTCTCAT	540
DB	701 CAATACCGCAAGCAGGGGGCGGTGGAACTGCCAAGGTGGGGGCTTTGGACTCTCTCAT	760
QY	541 CGATCCGCTGCGCTCTCTCTCATCTACAGTCTCTCAACAGGATTTACGAATACCAAGAT	600
DB	761 CGATCCGCTGCGCTCTCTCTCATCTACAGTCTCTCAACAGGATTTACGAATACCAAGAT	820
QY	601 GGGCGGCCCAAAATTTCCAAAGCCCTGTTAAGGTGGGAAGATGCAAAATGATGTCAAGA	660
DB	821 GGGCGGCCCAAAATTTCCAAAGCCCTGTTAAGGTGGGAAGATGCAAAATGATGTCAAGA	880
QY	661 ATGGCTTCTCATGGGAATCAAAATTTGTCAATTCAGCTTAAGAATGGGGGCAAAAGCTTACCA	720
DB	881 ATGGCTTCTCATGGGAATCAAAATTTGTCAATTCAGCTTAAGAATGGGGGCAAAAGCTTACCA	940
QY	721 GATACGTATCTCTTCAACACTGTGACAGAGATCACTGGAGCAAAATTCAGAAACAGGTT	780
DB	941 GATACGTATCTCTTCAACACTGTGACAGAGATCACTGGAGCAAAATTCAGAAACAGGTT	1000
QY	781 GTACTGTGCAATGGAACATCTGGAACAGCTGGGATGTGGCAGGGGTCCATGGAATGATGGC	840
DB	1001 GTACTGTGCAATGGAACATCTGGAACAGCTGGGATGTGGCAGGGGTCCATGGAATGATGGC	1060
QY	841 GGTGGAGGCTTTATATCAATGGAAGCACTCACTTAATTAAGATCTTTGGCTCGTCTCA	900
DB	1061 GGTGGAGGCTTTATATCAATGGAAGCACTCACTTAATTAAGATCTTTGGCTCGTCTCA	1120
QY	901 AAGAGGACTCTGGCGCTGGTCTCTGGAATGCAAGAAACAAGGTGAGATTGGTCTTTC	960
DB	1121 AAGAGGACTCTGGCGCTGGTCTCTGGAATGCAAGAAACAAGGTGAGATTGGTCTTTC	1180
QY	961 CAGATTTATCAGTTTACACAGGTAAATATTTTCCAACTACAGTCTGGGTGATGGAATCTGAC	1020
DB	1181 CAGATTTATCAGTTTACACAGGTAAATATTTTCCAACTACAGTCTGGGTGATGGAATCTGAC	1240

QY 1021 GCAGGAACCTTCTTACCACCTGGGCTGCATTCATCTGGCAGTGAAGGCGGCGCATC 1080  
DB 1241 GCAGGAACCTTCTTACCACCTGGGCTGCATTCATCTGGCAGTGAAGGCGGCGCATC 1300  
QY 1081 ATGAGAGAGGTTATGAGGCTGCTGCAGGCCCTCAATATCACTCAGTCTTGAGCCATGGA 1140  
DB 1301 ATGAGAGAGGTTATGAGGCTGCTGCAGGCCCTCAATATCACTCAGTCTTGAGCCATGGA 1360  
QY 1141 GAAGGAGACAGCATCACTTGTGATCCAGCTGAGGCTGAGCGACGCTTCTTGAT 1200  
DB 1361 GAAGGAGACAGCATCACTTGTGATCCAGCTGAGGCTGAGCGACGCTTCTTGAT 1420  
QY 1201 GACTTATACAGTATTTCTTCTTCATCACTCCACGAGACACATGATCATGAT 1260  
DB 1421 GACTTATACAGTATTTCTTCTTCATCACTCCACGAGACACATGATCATGAT 1480  
QY 1261 CCAAGAGAGATGAATGTTGCTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGACAGAC 1320  
DB 1481 CCAAGAGAGATGAATGTTGCTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGACAGAC 1540  
QY 1321 ATGAGAGAAATGCTGCTGATGCTC 1344  
DB 1541 ATGAGAGAAATGCTGCTGATGCTC 1564

RESULT 11  
AA298139  
ID AA298139 standard; cDNA; 1923 BP.  
AC AA298139;  
XX  
XX  
DT 11-MAY-2000 (first entry)  
XX  
DE Human signal peptide containing .protein HSPB-31 cDNA SEQ ID NO:165.  
XX  
XX Human; signal peptide-containing protein; HSPB; diagnosis; cancer;  
KM inflammation; cardiovascular disease; anticancer; anti-inflammation;  
KM antimicrobial; neuroprotective; cardiovascular; hepatotropic;  
KM antiasthmatic; gene therapy; cell proliferation; neurological disorder;  
KM reproductive disorder; developmental disorder; arteriosclerosis;  
KM cleftosis; psoriasis; acquired immune deficiency syndrome; anaemia;  
KM asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;  
KM Parkinson's disease; Huntington's disease; ovulatory defect;  
KM muscular dystrophy; ss.  
OS Homo sapiens.  
XX  
XX WO200000610-A2.  
XX  
XX PD 06-JAN-2000.  
XX  
XX PF 25-JUN-1999; 99WO-US14484.  
XX  
XX PR 26-JUN-1998; 98US-0090762.  
XX PR 31-JUL-1998; 98US-0094983.  
XX PR 01-OCT-1998; 98US-0102686.  
XX PR 11-DEC-1998; 98US-0112129.  
XX  
XX PA (INCY-) INCYTE PHARM INC.  
XX  
XX PI Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;  
PI Aherblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL,  
PI Bandman O;  
XX  
XX WPI; 2000-160673/14.  
XX DR P-PSDB; AAY87254.  
XX  
XX PT New human signal peptide-containing proteins useful in treatment,  
XX PT prevention and diagnosis of e.g. cancer, inflammation and  
XX PT cardiovascular disease  
XX  
XX PS Claim 9; Page 269-270; 327BP; English.

XX  
CC AA298109 to AA298242 encode AAY87224 to AAY87257 which represent the  
CC human signal peptide-containing proteins HSPB-1 to HSPB-134. HSPBs have  
CC anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatotropic,  
CC neuroprotective, cardiovascular and antiasthmatic activities, and can  
CC be used in gene therapy. HSPBs can be used to treat or prevent disorders  
CC associated with decreased activity or function of HSPB. Antagonists of  
CC HSPB are used to treat or prevent disorders associated with increased  
CC activity or function of HSPB. Such diseases include cell proliferation  
CC (including cancer), inflammation, cardiovascular, neurological,  
CC reproductive or developmental disorders, (e.g. arteriosclerosis,  
CC cleftosis, psoriasis, acquired immune deficiency syndrome, congestive  
CC asthma, Crohn's disease, microbial or other infections, anaemia or  
CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's  
CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSPB  
CC nucleic acids can be used for the recombinant production of HSPB, for  
CC detecting HSPB in standard hybridisation and amplification assays (for  
CC diagnosis and monitoring), in gene therapy, as antisense,  
CC triplex-forming or ribozyme therapeutics, for detecting related sequences  
CC or genetic variations, and for chromosomal mapping. HSPB are also used to  
CC raise specific antibodies (Ab) and to screen for agonists and  
CC antagonists (potential therapeutic agents). Ab are used to diagnose, or  
CC monitor, HSPB-related diseases (in usual immunoassays), as therapeutic  
CC antagonists, in competitive drug screens, and for purification of HSPB  
CC from natural sources.  
XX  
SQ Sequence 1923 BP; 538 A; 439 C; 471 G; 475 T; 0 other;  
Query Match 99.8%; Score 1340.8; DB 21; Length 1923;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AAGAAATGGCATCTCTTAAGAGACTTTTGAAGAAATTAAGAAATATGACCTGTGGA 60  
DB 251 AAGAAATGGCATCTCTTAAGAGACTTTTGAAGAAATTAAGAAATATGACCTGTGGA 310  
QY 61 GATGTTCTTAAGCAATCATCAACTGATGTTTATGTAAGCCGAGAACATCTCTAT 120  
DB 311 GATGTTCTTAAGCAATCATCAACTGATGTTTATGTAAGCCGAGAACATCTCTAT 370  
QY 121 GAGGATGTCACCTCTGTTGTAATCTGTTGACCCAGACTGATGCTCCAGAACCTTA 180  
DB 371 GAGGATGTCACCTCTGTTGTAATCTGTTGACCCAGACTGATGCTCCAGAACCTTA 430  
QY 181 GAAAAAGCCATCCAAATTAATGTAACCAAACTGACGAGAAATGAGGCTGAGAAATTCAC 240  
DB 431 GAAAAAGCCATCCAAATTAATGTAACCAAACTGACGAGAAATGAGGCTGAGAAATTCAC 490  
QY 241 CTGGAGCCAGTGAATATGCCCACTGGAGAGGGGAGAAATATGATGATGATGCTGAG 300  
DB 491 CTGGAGCCAGTGAATATGCCCACTGGAGAGGGGAGAAATATGATGATGATGCTGAG 550  
QY 301 CCAAGAAATTCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
DB 551 CCAAGAAATTCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 610  
QY 361 GGCATTACAGAGAAATGTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 420  
DB 611 GGCATTACAGAGAAATGTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 670  
QY 421 GAAGCAAGAGAGAAATGTTGTTTATTAACAACCTTACATCACTCAAGAGAGGATG 480  
DB 671 GAAGCAAGAGAGAAATGTTGTTTATTAACAACCTTACATCACTCAAGAGAGGATG 730  
QY 481 CAATACCGAAGCAGAGGGGCGGTGAGAGTCCAGAGGTGGGGCTTGGATCTCTCAT 540  
DB 731 CAATACCGAAGCAGAGGGGCGGTGAGAGTCCAGAGGTGGGGCTTGGATCTCTCAT 790  
QY 541 CGATCCGTGGCTCTCTTCTCAATCTACATGCTTCACAGAGATTCAGAAATACAGAT 600  
DB 791 CGATCCGTGGCTCTCTTCTCAATCTACATGCTTCACAGAGATTCAGAAATACAGAT 850  
QY 601 GCGGTGCCAAAATTCACAAGGCTGTATTAAGGTGGAATGAGAAATGATGTCAAGA 660

DB 851 GCGGCGCCCAAGATTCACAGCGCTGATTAAGGGAAGATGAGAATATGCAAGA 910  
 QY 661 ATGCTTCTCATGGAGTAAATTTGTCTATGAGCTAAAGATGGGGGCAAAAGCTACCCA 720  
 DB 911 ATGGCTTCTCATGGAGTAAATTTGTCTATGAGCTAAAGATGGGGGCAAAAGCTACCCA 970  
 QY 721 GATACGTATCTCTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAAAGGTT 780  
 DB 971 GATACGTATCTCTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAAAGGTT 1030  
 QY 781 GTACTGCTGAGTGAATCTGAGACAGCTGGAGATGTTGGGCAAGGTCATGATGATGCG 840  
 DB 1031 GTACTGCTGAGTGAATCTGAGACAGCTGGAGATGTTGGGCAAGGTCATGATGATGCG 1090  
 QY 841 GGTGAGGCTTTATATCATATGGAAGCACTCACTATTAAGATCTTGGGCTGCTGCA 900  
 DB 1091 GGTGAGGCTTTATATCATATGGAAGCACTCACTATTAAGATCTTGGGCTGCTGCA 1150  
 QY 901 AAGAGACTCTGCGGCTGCTGCTGAGCTGAGAGAAAGAGGAGGAGTGGCTTC 960  
 DB 1151 AAGAGACTCTGCGGCTGCTGCTGAGCTGAGAGAAAGAGGAGGAGTGGCTTC 1210  
 QY 961 CAGTATTATCAGTTACACAGGTAATATTTTCCACTACAGTCTGATGATGAGCTGAC 1020  
 DB 1211 CAGTATTATCAGTTACACAGGTAATATTTTCCACTACAGTCTGATGATGAGCTGAC 1270  
 QY 1021 GCAGGAACCTTTTACCACTGCGGCTGAGATTCATCTGAGCAATGAAAAGCCAGGCTTC 1080  
 DB 1271 GCAGGAACCTTTTACCACTGCGGCTGAGATTCATCTGAGCAATGAAAAGCCAGGCTTC 1330  
 QY 1081 ATGAGAGAGTTATGAGCTGCTGAGAGCCCTCAATATCATCTAGTCTGAGCCATGGA 1140  
 DB 1331 ATGAGAGAGTTATGAGCTGCTGAGAGCCCTCAATATCATCTAGTCTGAGCCATGGA 1390  
 QY 1141 GAAGGAGACAGATCACTTTTGTGATCCAAAGCTGAGTGCCTGAGCCAGTCTGAT 1200  
 DB 1391 GAAGGAGACAGATCACTTTTGTGATCCAAAGCTGAGTGCCTGAGCCAGTCTGAT 1450  
 QY 1201 GACTTATACAGTATTTCTTCCATCATCTCCACCGAGACACCATCATCTGATGAT 1260  
 DB 1451 GACTTATACAGTATTTCTTCCATCATCTCCACCGAGACACCATCATCTGATGAT 1510  
 QY 1261 CCAAGAGAGATGAATGTTGCTGCTGCTGTTGGGCTGTTTCTTATGTTGTCAGAC 1320  
 DB 1511 CCAAGAGAGATGAATGTTGCTGCTGCTGTTGGGCTGTTTCTTATGTTGTCAGAC 1570  
 QY 1321 ATGGAAGAAATGCTGCTGATGCTC 1344  
 DB 1571 ATGGAAGAAATGCTGCTGATGCTC 1594  
 RESULT 12  
 ACCS0603  
 ID ACCS0603 standard; cDNA; 2077 BP.  
 XX ACCS0603;  
 AC  
 XX 12-JUN-2003 (first entry)  
 DT  
 XX Human secreted protein coding sequence, SEQ ID 270.  
 DE  
 XX Cardiac; antiarrhythmic; antiarteriosclerotic; vasotropic; cyrostatic;  
 KW vulnerable; antiinflammatory; nootropic; neuroprotective;  
 KW antiParkinsonian; gene therapy; human; cardiovascular disorder;  
 KW gene; ss.  
 OS Homo sapiens.  
 XX  
 OS  
 PN W0200295010-A2.  
 XX  
 PD 28-NOV-2002.  
 XX

PF 19-MAR-2002; 2002WC-US09785.  
 XX  
 PR 21-MAR-2001; 2001US-277340P.  
 PR 19-JUN-2001; 2001US-306171P.  
 PR 13-NOV-2001; 2001US-331287P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen CA, Ruben SM;  
 DR WPI; 2003-129429/12.  
 XX  
 PT Novel human secreted proteins, useful for detecting, preventing,  
 PT diagnosing, prognosticating, treating and/or ameliorating  
 PT cardiovascular disorders such as arrhythmia -  
 XX  
 PS Claim 21; SEQ ID 270; 1881bp; English.  
 XX  
 CC The present invention relates to novel human secreted proteins  
 CC (ABR4763-ABR48145) and their coding sequences (ACCS0344-ACCS0856). The  
 CC proteins and their coding sequences are useful for the preparation of a  
 CC diagnostic or pharmaceutical composition for diagnosing or treating a  
 CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,  
 CC coronary arteriosclerosis, muscular disorders, reproductive disorders,  
 CC immune system disorders, pulmonary disorders, renal disorders,  
 CC gastrointestinal disorders, or cancerous diseases and conditions,  
 CC proliferative disorders and/or cancerous diseases and conditions, for  
 CC wound healing and epithelial cell proliferation, to treat inflammation or  
 CC infection, for treating thrombosis and arteriosclerosis, for treating or  
 CC preventing neural damage which occurs in neuronal disorders or  
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
 CC disease, to enhance bone and periodontal regeneration and aid in tissue  
 CC transplants or bone grafts, to prevent skin aging or hair loss, to  
 CC stimulate growth and differentiation of hematopoietic cells and bone  
 CC marrow cells when used in combination with other cytokines, to maintain  
 CC organs before transplantation or for supporting cell culture of primary  
 CC tissues, to increase or decrease differentiation or proliferation of  
 CC embryonic stem cells, or to modulate mammalian characteristics or  
 CC metabolism.  
 CC Note: The sequence data for this patent was published in electronic  
 CC format and is available from WIPO at  
 CC ftp.wipo.int/pub/1published\_pct\_sequences.  
 CC  
 XX  
 SQ Sequence 2077 BP; 597 A; 436 C; 482 G; 562 T; 0 other;  
 Query Match 99.8%; Score 1340.8; DB 25; Length 2077;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 AAGAAATGGCATCTCTAAGAGAGACTTTGAGAAATAAAGAAATAGCCAGCTGGA 60  
 DB 204 AAGAAATGGCATCTCTAAGAGAGACTTTGAGAAATAAAGAAATAGCCAGCTGGA 263  
 QY 61 GATGTTGCTAAGCAATCATCACTAGCTGTTATGTAAGCCGAAACAGATCTAT 120  
 DB 264 GATGTTGCTAAGCAATCATCACTAGCTGTTATGTAAGCCGAAACAGATCTAT 323  
 QY 121 GAGGATTTGGCACTTCTGTTGATCTGTTGAGACCCAGACTGAGTCTCAAGAACTTA 180  
 DB 324 GAGGATTTGGCACTTCTGTTGATCTGTTGAGACCCAGACTGAGTCTCAAGAACTTA 383  
 QY 181 GAAAAAGCATCCCAATTTATGTACCAAAACCTTGAGCAATGAGGCTGAGAAAGTTTAC 240  
 DB 384 GAAAAAGCATCCCAATTTATGTACCAAAACCTTGAGCAATGAGGCTGAGAAAGTTTAC 443  
 QY 241 CTGAGGCAAGTGAATATCCCACTGAGAGAGGAGAGAAATCACTGATGATCTGAG 300  
 DB 444 CTGAGGCAAGTGAATATCCCACTGAGAGAGGAGAGAAATCACTGATGATCTGAG 503  
 QY 301 CCAAGATTCATATAGATGAGCATCTGAGTCTTGGCAGCAGCATTTGGGACTCTCCAGAA 360  
 DB 504 CCAAGATTCATATAGATGAGCATCTGAGTCTTGGCAGCAGCATTTGGGACTCTCCAGAA 563

```

QY 361 GGCATTACAGCAGAGTTCTGTGTGATGACCTCTTTGATGATGACAGAGAGGCGCTCA 420
DB 564 GGCATTACAGCAGAGTTCTGTGTGATGACCTCTTTGATGATGACAGAGAGGCGCTCA 623
QY 421 GAAGCAAGAGGAGATTTGTTTATTAACCACTTAACATCACTACTCAAGACGGTG 480
DB 624 GAAGCAAGAGGAGATTTGTTTATTAACCACTTAACATCACTACTCAAGACGGTG 683
QY 481 CAATACGGAACGACGGGGGGGGTGGAAAGCTGCGCAAGGGGGGGGCTTTGGCATCTCTGAT 540
DB 684 CAATACGGAACGACGGGGGGGGTGGAAAGCTGCGCAAGGGGGGGGCTTTGGCATCTCTGAT 743
QY 541 CGATCCGTGGCTCTCTCTCATCTAGCTCTCAACAGGATTCAGAAATACAGGAT 600
DB 744 CGATCCGTGGCTCTCTCTCATCTAGCTCTCAACAGGATTCAGAAATACAGGAT 803
QY 601 GGGGTGCCCCAAATTCACAAGCTGTATTAAGGTGGAAGTGAAGATGATGATCAAGA 660
DB 804 GGGGTGCCCCAAATTCACAAGCTGTATTAAGGTGGAAGTGAAGATGATGATCAAGA 863
QY 661 ATGGCTCTCATGGGATCAAAATTTGATTCAGCTTAAGATGGGGGCAAAAGCACTACCA 720
DB 864 ATGGCTCTCATGGGATCAAAATTTGATTCAGCTTAAGATGGGGGCAAAAGCACTACCA 923
QY 721 GATACGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAATATCCAGAACAGGTT 780
DB 924 GATACGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAATATCCAGAACAGGTT 983
QY 781 GATACGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAATATCCAGAACAGGTT 840
DB 984 GATACGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAATATCCAGAACAGGTT 1043
QY 841 GGTGGAGCTTTATATATAGGAGAGCACTCTCACTTAATTAAGATCTTGGCTGCGTCCA 900
DB 1044 GGTGGAGCTTTATATATAGGAGAGCACTCTCACTTAATTAAGATCTTGGCTGCGTCCA 1103
QY 901 AAGAGACTCTGCGGCTGTGCTGTGACTGTGACAGAGAAACAAGGTGGAGTGGCTTC 960
DB 1104 AAGAGACTCTGCGGCTGTGCTGTGACTGTGACAGAGAAACAAGGTGGAGTGGCTTC 1163
QY 961 CAGTATTATCAGTTACACAAGATTAATTTTCCAACTACAGTCTGTGATGAGTCTGAC 1020
DB 1164 CAGTATTATCAGTTACACAAGATTAATTTTCCAACTACAGTCTGTGATGAGTCTGAC 1223
QY 1021 GCAGGAACCTTTTATACCACTGGGCTGCAATTAATCACTGGCAATGAAAAAGGCCAGCATC 1080
DB 1224 GCAGGAACCTTTTATACCACTGGGCTGCAATTAATCACTGGCAATGAAAAAGGCCAGCATC 1283
QY 1081 ATGAGAGAGGTTATGAGCTGTGAGGCTGCAATTAATCACTGGCAATGAAAAAGGCCAGCATC 1140
DB 1284 ATGAGAGAGGTTATGAGCTGTGAGGCTGCAATTAATCACTGGCAATGAAAAAGGCCAGCATC 1343
QY 1141 GAAGGAGCAGACATCAATTTTGAATCCAAAGCTGAGAGTCTGAGACCACTTATGAT 1200
DB 1344 GAAGGAGCAGACATCAATTTTGAATCCAAAGCTGAGAGTCTGAGACCACTTATGAT 1403
QY 1201 GACTTATTAACAAGTATTTCTTTCATATACCTCCACGAGACACCATATGATGAT 1260
DB 1404 GACTTATTAACAAGTATTTCTTTCATATACCTCCACGAGACACCATATGATGAT 1463
QY 1261 CCAAGAGCAGATGAATGTTGCTGTGCTGTTGGGGGTGTTTCTTATGTTGGACAGC 1320
DB 1464 CCAAGAGCAGATGAATGTTGCTGTGCTGTTGGGGGTGTTTCTTATGTTGGACAGC 1523
QY 1321 ATGAGAGAAATGCTGCTAGGTCC 1344
DB 1524 ATGAGAGAAATGCTGCTAGGTCC 1547

```

RESULT 13  
AB271331 standard; cdna; 2077 BP.  
XX

```

AC AB271331,
XX
XX 04-APR-2003 (first entry)
DE Secreted protein-encoding gene 142 cDNA clone HRA035, SEQ ID NO:152.
XX
XX Human, secreted protein; digestive disorder; gastrointestinal disorder;
XX mouth; oesophagus; stomach; small intestine; large intestine; liver;
XX biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
XX immune disorder; inflammation; infection; wound healing; drug screening;
XX chromosome identification; chromosome mapping; cytostatic; gene therapy;
XX anti-inflammatory; immunosuppressive; vulnertary; chromosome 8q22.2;
XX gene; ss.
XX Homo sapiens.
XX
XX W0200276488-A1.
XX
XX 03-OCT-2002.
XX
XX 19-MAR-2002; 2002MO-US08276.
XX
XX 21-MAR-2001; 2001US-277340P.
XX
XX 19-JUL-2001; 2001US-306171P.
XX
XX 13-NOV-2001; 2001US-331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-029900/02.
XX
XX P-PSDB; ABR00152.
XX
XX New human secreted proteins and nucleic acids, useful for detecting,
XX preventing, diagnosing, prognosticating, treating and/or ameliorating
XX e.g. gastrointestinal diseases and disorders, or cancers -
XX
XX Claim 21; Page 841-842; 1216pp; English.
XX
XX AB271190-AB271478 represent cDNAs corresponding to 178 human secreted
XX protein genes, and ABR00011-ABP00299 represent the proteins they encode.
XX AB271479-AB271540 represent human secreted protein genomic fragments. The
XX invention also encompasses antibodies specific for the secreted proteins,
XX the use of the secreted proteins in drug screening, and recombinant
XX vectors and host cells comprising a nucleic acid of the invention. The
XX secreted proteins, nucleic acids encoding them, antibodies or antibody
XX fragments specific for the secreted proteins, and modulators of protein
XX activity are useful for diagnosing, treating, ameliorating or preventing
XX digestive disorders. Such conditions include disorders of the mouth,
XX oesophagus, stomach, small intestine, large intestine, liver, biliary
XX tract and pancreas, and include cancers of these organs and tissues. The
XX secreted proteins and their nucleic acids may also be used in the
XX treatment of immune disorders, inflammation, infection,
XX hyperproliferative disorders, and to promote wound healing. Nucleic acids
XX of the invention may be used for chromosome identification, chromosome
XX mapping, in gene therapy, for identifying individuals from minute
XX biological samples, as hybridisation probes, and as molecular weight
XX markers. The present sequence represents a human secreted protein-
XX encoding cDNA clone of the invention.
XX
XX Sequence 2077 BP; 597 A; 436 C; 482 G; 562 T; 0 other;
XX
XX Query Match 99.8%; Score 1340.8; DB 25; Length 2077;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AAGATGGCATCTCTTAAGAGACTTTGAAGAATAAAGAAATAGCAGCTGGA 60
DB 204 AAGATGGCATCTCTTAAGAGACTTTGAAGAATAAAGAAATAGCAGCTGGA 263
QY 61 GATGTTCTAAGCAATCATCACTAGCTGTTTATGTAAGCCAGAAACAGATCTAT 120
DB 264 GATGTTCTAAGCAATCATCACTAGCTGTTTATGTAAGCCAGAAACAGATCTAT 323

```



```

QY 121 GAGGATTTGGCACTTCTGGTGTATCTGTTGGACCCAGACTGAGTGGCTCCAGAACTTA 180
D 324 GAGGATTTGGCACTTCTGGTGTATCTGTTGGACCCAGACTGAGTGGCTCCAGAACTTA 383
QY 181 GAAAAAGCCATCCAAATTTATGTATCCAAAACCTGACAGAAATGGCTGAGAAAGTTGAC 240
D 384 GAAAAAGCCATCCAAATTTATGTATCCAAAACCTGACAGAAATGGCTGAGAAAGTTGAC 443
QY 241 CTGAGAGCCAGTGAAGTATCCCACTGGAGAGGGGAGAAATGACGCTGATGCTGGAG 300
D 444 CTGAGAGCCAGTGAAGTATCCCACTGGAGAGGGGAGAAATGACGCTGATGCTGGAG 503
QY 301 CCAAGAAATTCATTAAGATAGCCATCTGGGCTTTGGACACACATTTGGGACTCTCCAGAA 360
D 504 CCAAGAAATTCATTAAGATAGCCATCTGGGCTTTGGACACACATTTGGGACTCTCCAGAA 563
QY 361 GGCATTACAGCAAGAAATTTCTGGTGTGAACCTTTTGTGATGAACCTGACAGAAAGGCTTCA 420
D 564 GGCATTACAGCAAGAAATTTCTGGTGTGAACCTTTTGTGATGAACCTGACAGAAAGGCTTCA 623
QY 421 GAACCAAGAGGAAATTTGTTTATTAACCACTTACATCACTACTCAAGAGCGGTG 480
D 624 GAACCAAGAGGAAATTTGTTTATTAACCACTTACATCACTACTCAAGAGCGGTG 683
QY 481 CAATACCGAAGCAAGGGGGCGGTGGAGCTGCAAGGTGGGGCTTTGGCATCTCTCAT 540
D 684 CAATACCGAAGCAAGGGGGCGGTGGAGCTGCAAGGTGGGGCTTTGGCATCTCTCAT 743
QY 541 CGATCCGTGGCTCTCTCTCATCTACAGTCTCAACAGGATTTACAGAAATACAGGAT 600
D 744 CGATCCGTGGCTCTCTCTCATCTACAGTCTCTCAACAGGATTTACAGAAATACAGGAT 803
QY 601 GGCCGCCCCAAATTTCCAAAGCCCTGTATTAACGTGGAGAAATGCAGAAATGATGTCAGA 660
D 804 GGCCGCCCCAAATTTCCAAAGCCCTGTATTAACGTGGAGAAATGCAGAAATGATGTCAGA 863
QY 661 ATGGCTTCTCATGGGATCAAAATTTGTCAATTCAGTCAAGTGGGGGCAAAAGACTTACCCA 720
D 864 ATGGCTTCTCATGGGATCAAAATTTGTCAATTCAGTCAAGTGGGGGCAAAAGACTTACCCA 923
QY 721 GATACTGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAAAGGTT 780
D 924 GATACTGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAAAGGTT 983
QY 781 GTACTGTCAGTGGACATCTGGAACAGCTGGGATGTTGGGCAAGGTCATGATGATGGC 840
D 984 GTACTGTCAGTGGACATCTGGAACAGCTGGGATGTTGGGCAAGGTCATGATGATGGC 1043
QY 841 GGTGGAGCCTTTATATCATGAGGAAGCACTCTCACTTATTAAGATCTTGGGCTGGCTCA 900
D 1044 GGTGGAGCCTTTATATCATGAGGAAGCACTCTCACTTATTAAGATCTTGGGCTGGCTCA 1103
QY 901 AAGAGACTCTGCGGCTGGTGTCTCTGACTGACAGAAAGCAAGGTGAGTGGTCCCTTC 960
D 1104 AAGAGACTCTGCGGCTGGTGTCTCTGACTGACAGAAAGCAAGGTGAGTGGTCCCTTC 1163
QY 961 CAGATTTATCAGTTTACAAAGTAAATTTTCCAACTACAGTCTGGTATGAGAGTCTGAC 1020
D 1164 CAGATTTATCAGTTTACAAAGTAAATTTTCCAACTACAGTCTGGTATGAGAGTCTGAC 1223
QY 1021 GCAGGAACCTTTACCACTGGGCTGCAATTCATGSCAGTGAAGAAAGCCAGGGCCATC 1080
D 1224 GCAGGAACCTTTACCACTGGGCTGCAATTCATGSCAGTGAAGAAAGCCAGGGCCATC 1283
QY 1081 ATGAGAGAGTTATGAGCTCTGCAAGCCCTCATATATCACTCAAGTCTTGAGCCATGGA 1140
D 1284 ATGAGAGAGTTATGAGCTCTGCAAGCCCTCATATATCACTCAAGTCTTGAGCCATGGA 1343
QY 1141 GAAAGGACAGACATCACTTTTGGATCCAAGCTGAGTGGCTGGAAGCCAGTCTTACTTGAT 1200
D 1344 GAAAGGACAGACATCACTTTTGGATCCAAGCTGAGTGGCTGGAAGCCAGTCTTACTTGAT 1403

```

```

QY 1201 GACTTATACAGATATTTCTTCTTCATCACTCTCCACGAGACACATGATGTCAGAT 1260
D 1404 GACTTATACAGATATTTCTTCTTCATCACTCTCCACGAGACACATGATGTCAGAT 1463
QY 1261 CCAAGCAGATGAATGTTGCTGCTGCTGTTGGCTGTGTTGTTCTTATGTTTGCAGAC 1320
D 1464 CCAAGCAGATGAATGTTGCTGCTGCTGTTGGCTGTGTTGTTCTTATGTTTGCAGAC 1523
QY 1321 ATGGAAGAAATGCTGCTAGGTC 1344
D 1524 ATGGAAGAAATGCTGCTAGGTC 1547

RESULT 14
ID AAK94491 standard; cDNA; 1860 BP.
XX
AC AAK94491;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA, SEQ ID NO: 3328.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
EN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI; 2001-524255/58.
DR P-PSDB; AAM93559.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 8; SEQ ID NO 3328; 1380bp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 1860 BP; 515 A; 423 C; 451 G; 471 T; 0 other;

Query Match 99.4%; Score 1336; DB 22; Length 1860;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1339; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAGATGGCATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGGA 60
D 201 AAGATGGCATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGGA 260

```



QY 61 GATGTTGCTAAGCAATCATCAACCTAGCTGTTATATGTTAAAGCCAGACAAGATCTTAT 120  
 Db 261 GATGTTGCTAAGCAATCATCAACCTAGCTGTTATATGTTAAAGCCAGACAAGATCTTAT 320  
 QY 121 GAGGATTTGGACATTTCTGTTGATATCTGTTGAGCCCAAGCTGAGTGGCTCCAGAACCTA 180  
 Db 321 GAGGATTTGGACATTTCTGTTGATATCTGTTGAGCCCAAGCTGAGTGGCTCCAGAACCTA 380  
 QY 181 GAAAAAGCCATCCAAATATATGATACCAAAAACCTGACGAAGATGGGCTGGAGAAAGTTTAC 240  
 Db 381 GAAAAAGCCATCCAAATATATGATACCAAAAACCTGACGAAGATGGGCTGGAGAAAGTTTAC 440  
 QY 241 CTGAGCCAGTGAAGATACTCCCACTGGAGAGGGAGAGAAATCACTGTATGCTGGAG 300  
 Db 441 CTGAGCCAGTGAAGATACTCCCACTGGAGAGGGAGAGAAATCACTGTATGCTGGAG 500  
 QY 301 CCAAGAAATTCATTAAGATATGCCATCTCTGGCTCTTGGACGACGATTTGGATCTCTCAGAA 360  
 Db 501 CCAAGAAATTCATTAAGATATGCCATCTCTGGCTCTTGGACGACGATTTGGATCTCTCAGAA 560  
 QY 361 GGCATTTACAGAGAGATCTGCTGGTGAATCTTTGATGAATGTCAGAGAAAGGCTTCA 420  
 Db 561 GGCATTTACAGAGAGATCTGCTGGTGAATCTTTGATGAATGTCAGAGAAAGGCTTCA 620  
 QY 421 GAAGCAAGAGGAGATTTGTTTATATACCAACCTTACATCACTACTCAAGAGCGTGC 480  
 Db 621 GAAGCAAGAGGAGATTTGTTTATATACCAACCTTACATCACTACTCAAGAGCGTGC 680  
 QY 481 CAATACCGAAGCGAGAGGGGCGGTGAGCTGCCAAGGTGGGGGCTTTGGCATCTCTCAT 540  
 Db 681 CAATACCGAAGCGAGAGGGGCGGTGAGCTGCCAAGGTGGGGGCTTTGGCATCTCTCAT 740  
 QY 541 CGATCCGCGGCTCTCTCTCATCTACAGTCTCTCAAGAGATTCAGAGATTCACAGAT 600  
 Db 741 CGATCCGCGGCTCTCTCTCATCTACAGTCTCTCAAGAGATTCAGAGATTCACAGAT 800  
 QY 601 GGCCTGCCCCAAATTTCCAAAGCCTGTATTAAGGTGAGAAATGCAAGATATGTCAGAA 660  
 Db 801 GGCCTGCCCCAAAGATTTCCAAAGCCTGTATTAAGGTGAGAAATGCAAGATATGTCAGAA 860  
 QY 661 ATGCTTCTCATTTGGATCAAAATTTGTCATTCAGCTTAAGATGGGGGCAAAAGCCTACCA 720  
 Db 861 ATGCTTCTCATTTGGATCAAAATTTGTCATTCAGCTTAAGATGGGGGCAAAAGCCTACCA 920  
 QY 721 GATCTGATTCCTTCAACACTGTGAGAGAGATTCCTGGAGCAAAATATCCAGAACAGTT 780  
 Db 921 GATCTGATTCCTTCAACACTGTGAGAGAGATTCCTGGAGCAAAATATCCAGAACAGTT 980  
 QY 781 GTACTGCTGATGACATCTGGAAGCTGGATGTTGGGCAAGGCTCCATGATGATGATGATG 840  
 Db 981 GTACTGCTGATGACATCTGGAAGCTGGATGTTGGGCAAGGCTCCATGATGATGATGATG 1040  
 QY 841 GGTGAGGCTTTATATCATGGAAGCACTCTCACTTATTAAGATCTTTGGCTGGCTGCA 900  
 Db 1041 GGTGAGGCTTTATATCATGGAAGCACTCTCACTTATTAAGATCTTTGGCTGGCTGCA 1100  
 QY 901 AAGAGGACTCTGGCGCTGCTGCTGGAATGCAAGAAACAAGTGGAGTGGTGGCTTC 960  
 Db 1101 AAGAGGACTCTGGCGCTGCTGCTGGAATGCAAGAAACAAGTGGAGTGGTGGCTTC 1160  
 QY 961 CAGATTTATCAGTTACACAAGATTAATTTTCCAACTACAGTCTGATGATGATGATGATG 1020  
 Db 1161 CAGATTTATCAGTTACACAAGATTAATTTTCCAACTACAGTCTGATGATGATGATGATG 1220  
 QY 1021 GCAGGAACCTTTTACCCAGTGGGCTGCAATTCATCTGGCACTGAAAGGCAAGGCTATC 1080  
 Db 1221 GCAGGAACCTTTTACCCAGTGGGCTGCAATTCATCTGGCACTGAAAGGCAAGGCTATC 1280  
 QY 1081 ATGAGAGAGGTTATGAGGCTGCTGAGGCTCCCAATATCACTCAGTCTTGAAGCATGAG 1140  
 Db 1281 ATGAGAGAGGTTATGAGGCTGCTGAGGCTCCCAATATCACTCAGTCTTGAAGCATGAG 1340  
 QY 1141 GAAAGGAGAGAGATCACTTTTGGATCAAGCTGAGTGGCTGGAGGCAAGTCTTATGAT 1200

Db 1341 GAAAGGAGAGAGATCACTTTTGGATCAAGCTGAGTGGCTGGAGGCAAGTCTTATGAT 1400  
 QY 1201 GACTTATACAGATTTTCTTTTCCATCATCTCCACGAGACACCAATGATCATGATGAT 1260  
 Db 1401 GACTTATACAGATTTTCTTTTCCATCATCTCCACGAGAGACACCAATGATCATGATGAT 1460  
 QY 1261 CCAAGGAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320  
 Db 1461 CCAAGGAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1520  
 QY 1321 ATGAGAGAAATGCTGCTAGGCTC 1344  
 Db 1521 ATGAGAGAAATGCTGCTAGGCTC 1544  
 RESULT 15  
 AAH99703  
 ID AAH99703 standard; cDNA, 1895 BP.  
 AC AAH99703;  
 XX  
 DT 16-OCT-2001 (first entry)  
 XX  
 DE Human protein encoding cDNA sequence SEQ ID NO:538.  
 XX  
 KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
 KW antihistaminic; antihistaminic; antihistaminic; antihistaminic;  
 KW antihistaminic; antihistaminic; antihistaminic; antihistaminic;  
 KW anti-HIV; fungicide; antitumor; cardiovascular; antineoplastic; anaemia;  
 KW antilegionella; haemostatic; vulnery; antilegionella; antilegionella;  
 KW dermatological; antiallergic; antiallergic; antiallergic; antiallergic;  
 KW neuroprotection; antidepressant; neurotropic; antiparkinsonian; infection;  
 KW immunomodulator; gene therapy; antitumor; antitumor; antitumor;  
 KW antihistaminic; antihistaminic; antihistaminic; antihistaminic;  
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
 KW chromocytopenia; osteoporosis; severe combined immunodeficiency;  
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
 KW neurological disorder; ss.  
 OS Homo sapiens.  
 XX  
 PN MO200153455-A2.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 22-DEC-2000; 2000WO-US35017.  
 XX  
 PR 23-DEC-1999; 99US-0471275.  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI; 2001-457603/49.  
 DR P-PSDB; AAM25762.  
 XX  
 PT Isolated human polynucleotides encoding polypeptides, useful for the  
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -  
 XX  
 XX Claim 1; Page 591; 1217p; English.  
 PS  
 CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
 CC AAM25963. The proteins can have activities based on the tissues and  
 CC cells they are expressed in, such as: antihistaminic; antihistaminic;  
 CC antihistaminic; immunosuppressive; antihistaminic; endocrine; cardiant;  
 CC central nervous system; virucide; anti-HIV; fungicide; antitumor;  
 CC cardiovascular; antineoplastic; antilegionella; haemostatic; vulnery;  
 CC antilegionella; antineoplastic; antilegionella; antilegionella;  
 CC antilegionella; antineoplastic; antilegionella; antilegionella;

CC antidiabetic; cyostatic; neuroprotective; antidepressant; nootropic;  
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
 CC encoding them can be used in gene therapy, antisense therapy and vaccine  
 CC production. The proteins and polynucleotides are useful for screening for  
 CC agonists or antagonists of a protein and for the treatment and diagnosis  
 CC of disorders associated with the activity of a protein e.g. inflammation,  
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,  
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
 CC neurological disorders.

XX Sequence 1895 BP; 530 A; 439 C; 450 G; 476 T; 0 other;

Query Match 95.5%; Score 1283.6; DB 22; Length 1895;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1335; Conservative 0; Mismatches 9; Indels 6; Gaps 4;

QY 1 AAGATGGCATCTCTAAGAGGCTTTTGAAGAAATGAAGAAATGCCAGCTGTGA 60  
 DB 198 AAGATGGCATCTCTAAGAGGCTTTTGAAGAAATGAAGAAATGCCAGCTGTGA 257  
 QY 61 GATGTTGCTAAAGCAATCAACCTAGCTTTTATGTTAAAGCCAGAACGATCTAT 120  
 DB 258 GATGTTGCTAAAGCAATCAACCTAGCTTTTATGTTAAAGCCAGAACGATCTAT 317  
 QY 121 GAGCGATTGGCACTTCTGGTTGATACCTGTGGACCCAGACTGAGTGGCTCCAGAACCTTA 180  
 DB 318 GAGCGATTGGCACTTCTGGTTGATACCTGTGGACCCAGACTGAGTGGCTCCAGAACCTTA 377  
 QY 181 GAAAAAGCATCCAAATATGTAACCAAACTGCGAGCAAGTGGCTGGAAGAAATTTAC 240  
 DB 378 GAAAAAGCATCCAAATATGTAACCAAACTGCGAGCAAGTGGCTGGAAGAAATTTAC 437  
 QY 241 CTGAGCCAGTGAAGATACCCCACTGGAGAGGGGAGAAATCACTGTGTGTGGAG 300  
 DB 438 CTGAGCCAGTGAAGATACCCCACTGGAGAGGGGAGAAATCACTGTGTGTGGAG 497  
 QY 301 CCAGAAATTCATAGATAGCCATCTGGGCTCTGGCAGCAGCATTTGGAGCTTCTCCAGAA 360  
 DB 498 CCAGAAATTCATAGATAGCCATCTGGGCTCTGGCAGCAGCATTTGGAGCTTCTCCAGAA 557  
 QY 361 GGCATTACAGCAGAGTCTGGTGGTGAACCTTCTTGATGAACCTGAGAAAGGGCTCA 420  
 DB 558 GGCATTACAGCAGAGTCTGGTGGTGAACCTTCTTGATGAACCTGAGAAAGGGCTCA 617  
 QY 421 GAAGCAAGAGGAGATTTGTTTATACCAACCTTACATCACTCAAGAGCGGTG 480  
 DB 618 GAAGCAAGAGGAGATTTGTTTATACCAACCTTACATCACTCAAGAGCGGTG 677  
 QY 481 CAATGCCAAGCGCAGGGGGGCTGGAGCTGCAAGGTGGGGGCTTTGGCATTTCTCAAT 540  
 DB 678 CAATGCCAAGCGCAGGGGGGCTGGAGCTGCAAGGTGGGGGCTTTGGCATTTCTCAAT 737  
 QY 541 CGATCGGCGCTCTCTTCTCACTCAAGTCTCTCAACAGGTAATCAGAAATACAGGAT 600  
 DB 738 CGATCGGCGCTCTCTTCTCACTCAAGTCTCTCAACAGGTAATCAGAAATACAGGAT 797  
 QY 601 GGCCTGCCCAAAATTCACACAGCTGTATACGTTGAAGATGCAAGATGTCACAGA 660  
 DB 798 GGCCTGCCCAAAATTCACACAGCTGTATACGTTGAAGATGCAAGATGTCACAGA 857  
 QY 661 ATGGCTTCTCATGGGATCAAAATTTGTCATTCAGTTAAAGATGGGGGCAAGACCTACCA 720  
 DB 858 ATGGCTTCTCATGGGATCAAAATTTGTCATTCAGTTAAAGATGGGGGCAAGACCTACCA 917  
 QY 721 GATACGTGATCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCGAACAGTT 780  
 DB 918 GATACGTGATCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCGAACAGTT 977

QY 781 GTACTGTCAGTGAAGATCTTGACAGCTGGGATTTGGGAGGGGTGCCATGATGATGCC 840  
 DB 978 GTACTGTCAGTGAAGATCTTGACAGCTGGGATTTGGGAGGGGTGCCATGATGATGCC 1037  
 QY 841 GGTGAGCCTTTATATATCATGAGGAAGCACTTCACTTATTAAGATCTTGGGCTGCTCA 900  
 DB 1038 GGTGAGCCTTTATATATCATGAGGAAGCACTTCACTTATTAAGATCTTGGGCTGCTCA 1097  
 QY 901 AAGAGACTGCGCGGCTGGTCTCTGACATGCGAGAAACAAAGGTGAGTTGCTGCTTC 960  
 DB 1098 AAGAGACTGCGCGGCTGGTCTCTGACATGCGAGAAACAAAGGTGAGTTGCTGCTTC 1157  
 QY 961 CAGATTATCAGTATACCAAGATTAATATTTCCAACTACAGTCTGATGATGAGTGCAC 1020  
 DB 1158 CAGATTATCAGTATACCAAGATTAATATTTCCAACTACAGTCTGATGATGAGTGCAC 1217  
 QY 1021 GCAGAACTTTCTTACCCTGCTGCAATTCATGCGAGTGAAGAAAGCCAGGGCCATC 1080  
 DB 1218 GCAGAACTTTCTTACCCTGCTGCAATTCATGCGAGTGAAGAAAGCCAGGGCCATC 1277  
 QY 1081 ATGAGAGGTTATAGACCTGCTGCAATTCATGCGAGTGAAGAAAGCCAGGGCCATC 1140  
 DB 1278 ATGAGAGGTTATAGACCTGCTGCAATTCATGCGAGTGAAGAAAGCCAGGGCCATC 1337  
 QY 1141 GAAGGACAGACATCACTTTTGGATCCAGTGAAGTGCCTGAGCCAGTCTTACTTAT 1200  
 DB 1338 GAAGGACAGACATCACTTTTGGATCCAGTGAAGTGCCTGAGCCAGTCTTACTTAT 1397  
 QY 1201 GACTTATACAGATATTTCTTTCATCACTCCACGAGACACCATGACTGTC--ATG 1257  
 DB 1398 GACTTATACAGATATTTCTTTCATCACTCCACGAGACACCATGACTGTC--ATG 1457  
 QY 1258 GATCCAAAGAGATGATG--TTGCTGCTGCTG--TTTGGGCTGTTTCTTATGT--TGT 1314  
 DB 1458 ATCCAAAGAGATGATGTTGCTGCTGCTGTTTGGGCTGTTTCTTATGTGTTT 1517  
 QY 1315 GCAGACATGAGAAAGATGCTGCTAGGCTC 1344  
 DB 1518 GCAGACATGAGAAAGATGCTGCTAGGCTC 1547

Search completed: December 22, 2003, 15:10:06  
 Job time : 328.93 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:22:34 ; Search time 2630.64 Seconds  
(without alignments)  
12417.225 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_171\_1514  
Perfect score: 1344  
Sequence: 1 AAGAAATGCTCTTAAG.....AAGAAATGCTCTTAAGTCC 1344

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estln:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rnd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrt:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query length	ID	Description
1	1088	81.0	1868	11 BC010977 Mus muscu
2	1086.4	80.8	1739	11 AK032972 Mus muscu
3	1084.8	80.7	1805	11 AK075686 Mus muscu
4	886.8	66.0	1024	11 BC012019 Homo sapi

5	833.8	62.0	1050	13 BX439467
6	771.8	57.4	1038	13 BX355940
7	740.6	55.1	866	13 BU146905
8	740	55.1	902	10 BG751497
9	726.6	54.1	996	13 BX355939
10	724.8	53.9	909	13 BQ878966
11	719.2	52.3	739	14 CB958693
12	702.4	52.3	1126	13 BX360507
13	694.2	51.7	919	13 BX355995
14	682.4	50.8	813	14 CD519131
15	681.8	50.7	708	14 CD366462
16	672.4	50.0	1012	13 BX416896
17	664.6	49.4	998	13 BX460463
18	653.2	48.6	1020	13 BX439466
19	636.8	47.4	653	10 BE218907
20	630.4	46.9	650	10 BG432644
21	626.2	46.6	804	14 CD352906
22	625.6	46.5	677	10 BE670671
23	614.8	45.7	1073	13 BX360506
24	612.2	45.6	889	10 BE906771
25	611.4	45.5	922	12 BG761741
26	604.6	45.0	945	12 BI909780
27	604	44.9	926	13 BQ938234
28	599.6	44.6	805	10 BG563740
29	597.4	44.4	3151	11 BC017373
30	594.8	44.3	1081	13 BQ072892
31	591.6	44.0	1081	12 BI754143
32	591.2	44.0	1003	13 BX416895
33	583.8	43.4	701	12 BI771109
34	576.8	42.9	751	14 CD558931
35	570.4	42.4	786	10 BG430966
36	569.8	42.4	783	12 BI754468
37	568.4	42.3	929	13 BX355994
38	560.8	41.7	822	12 BI911772
39	560	41.7	944	12 BI819273
40	554.2	41.0	879	13 BQ218838
41	551	41.0	937	14 CD515740
42	542.4	40.4	544	10 AW952474
43	534.6	39.8	794	14 CD356404
44	534.4	39.8	774	14 CB308109
45	530.2	39.4	727	12 BG765433

## ALIGNMENTS

RESULT 1  
BC010977  
LOCUS  
DEFINITION Mus musculus, similar to plasma glutamate carboxypeptidase, clone  
IMAGE:4018296, mRNA.  
ACCESSION BC010977  
VERSION BC010977.1 GI:15012137  
KEYWORDS HTC.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (23-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbioology.org>

contact: amadan@systembiology.org  
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILM, at: <http://image.llnl.gov>  
Series: IRAX Plate: 18 Row: 9 Column: 4  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9055233  
This clone has the following problem: retained intron.

## FEATURES

source

Location/Qualifiers

1..1868  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6 J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4018296"  
/issue\_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."  
/clone\_id="NCI CGAP\_Lu29"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"

BASE COUNT 505 a 418 c 460 g 485 t

ORIGIN

Query Match 81.0%; Score 1088; DB 11; Length 1868;  
Best Local Similarity 88.1%; Pred. No. 0;  
Matches 1184; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

1 AAGATGCGATCTCTAAGAGAGACTTTGAAGAAATTAAGAAATATGAGCCAGCTGTGA 60  
267 AAGATGCGATCTCTAAGAGAGACTTTGAAGAAATTAAGAAATATGAGCCAGCTGTGA 326  
61 GATGTTGCTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCAGAAAGATCTTAT 120  
327 GATGTTGCTAAAGCAATCATCACTAGCTGTTATGTTAAAGCCAGAAAGATCTTAT 386  
121 GAGCGATTGCGATCTCTGTTGATATCTGTGGAAGCCAGCTGAGTGTCCAGAACTTA 180  
387 GAGCGATTGCGATCTCTGTTGATATCTGTGGAAGCCAGCTGAGTGTCCAGAACTTA 446  
181 GAAAAAGCCATCAATTAATGTAACCAAACTGCGAGCAAGATGGCTGGAAGATTTGAC 240  
447 GAGAAAGCTATCCAAATCATGTAACCAAACTGCGAGCAAGATGGCTGGAAGATTTGAC 506  
241 CTGAGCGAGTGAAGTAACCCCACTGGGAGAGGAGGAAGAAATCAGTGTGTGAG 300  
507 CTGAGCGAGTGAAGTAACCCCACTGGGAGAGGAGGAAGAAATCAGTGTGTGAG 566  
301 CCAAGAAATTCATAGATAGCCATCTCTGGGTCTTGGCAGCAGATTTGGATCTCTCAGA 360  
567 CCAAGAAATTCATAGATAGCCATCTCTGGGTCTTGGCAGCAGATTTGGATCTCTCAGA 626  
361 GGCATTACAGCAGAGTCTGTGGTGTGACCTCTTTCAGTGAATCTGAGAGAGGCTTCA 420  
627 GGCATTACAGCAGAGTCTGTGGTGTGACCTCTTTCAGTGAATCTGAGAGAGGCTTCA 686  
421 GAAGCAAGAGGAGATGTTGTTTATTAACCAACTTATCATCACTACTCAAGAGGCTG 480  
687 GAAGCAAGAGGAGATGTTGTTTATTAACCAACTTATCATCACTACTCAAGAGGCTG 746  
481 CAAATCCGAAAGCGAGGCGGCTGTGAAGTGTGCAAGGTGGGGCTTTGGCATCTTCTATT 540  
747 CAAATCCGAAAGCGAGGCGGCTGTGAAGTGTGCAAGGTGGGGCTTTGGCATCTTCTATT 806  
541 CGATCGTGGGCTCTCTTCCATCTACAGTCTCTCAACAGGATTCAGGAATTCAGAGAT 600  
807 CAAATGAGTGGCTCTCTTCCATCTACAGTCTCTCAACAGGATTCAGGAATTCAGAGAT 866  
601 GGCCTGCCCAAAATTCAGAGCGCTGTATTCAGGTGAAGATGAGGAATGATGTCAAGA 660  
867 GGTGTGCCCAAGATTCAGAGCGCTGTATTCAGGTGAAGATGAGGAATGATGTGTCTGA 926

661 ATGGCTTCTCATGGGATCAAAATTTGTCATTGAGTAAAGATGGGGGCAAGAACTTACCA 720  
927 ATGGCTTCTCATGGGATCAAAATTTGTCATTGAGTAAAGATGGGGGCAAGAACTTACCA 986  
721 GATACGATTTCTTCAACACTGTGAGAGATCACTGGGAGCAAAATTCAGAAACAGATT 780  
987 GATACGATTTCTTCAACACTGTGAGAGATCACTGGGAGCAAAATTCAGAAACAGATT 1046  
781 GTACTGTGAGTGAACATCTGAGACGCTGGATTTGGGCGAGGTCCTCATGATGATGCC 840  
1047 GTCTGTGAGTGAACATCTGAGACGCTGGATTTGGGCGAGGTCCTCATGATGATGAT 1106  
841 GGTGAGCGCTTTATATATGAGGAAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCA 900  
1107 GTGAGAGCTTCAATATATGAGGAAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCA 1166  
901 AAGAGACTCTGCGGCTGTGCTCTGAGACTGCAAGAAAGAGTGAAGTGTGCTTTC 960  
1167 AAGAGACTCTGCGGCTGTGCTCTGAGACTGCAAGAAAGAGTGAAGTGTGCTTTC 1226  
961 CAGTATTATCACTTACCAAGATTAATTTTCCATCACTGCTGTGATGAGTGTGAC 1020  
1227 CAGTATTATCACTTACCAAGATTAATTTTCCATCACTGCTGTGATGAGTGTGAC 1286  
1021 GCAGAACTCTTACCAAGATTAATTTTCCATCACTGCTGTGATGAGTGTGAC 1080  
1287 TCAGAACTCTTACCAAGATTAATTTTCCATCACTGCTGTGATGAGTGTGAC 1346  
1081 ATGAGAGAGGTTATGAGCTGCTGCAAGCCCTCAATATCACTGCTGTGAGTGA 1140  
1347 ATGAGAGAGGTTATGAGCTGCTGCAAGCCCTCAATATCACTGCTGTGAGTGA 1406  
1141 GAAGGACAGATCAATCACTTGTGATCAAGCTGAGTGTGCTGAGCCAGTCTTATGAT 1200  
1407 GAAGGACAGATCAATCACTTGTGATCAAGCTGAGTGTGCTGAGCCAGTCTTATGAT 1466  
1201 GACTTATCAAGATTTCTTCTTCCATCACTGCTGCAAGCCCTCAATATCACTGCTGAGT 1260  
1467 GACTTATCAAGATTTCTTCTTCCATCACTGCTGCAAGCCCTCAATATCACTGCTGAGT 1526  
1261 CCAAGCAAGATGATGTTGCTGCTGCTGTTGGGCTGTGTTCTTATGTTGTCAGAC 1320  
1527 CCAAGCAAGATGATGTTGCTGCTGCTGTTGGGCTGTGTTCTTATGTTGTCAGAC 1586  
1321 ATGGAAGAAATGCTGCTAGTCC 1344  
1587 ATGGAAGAAATGCTGCTAGTCC 1610

RESULT 2  
AK032972  
LOCUS  
DEFINITION  
MUS MUSCULUS 12 DAYS EMBRYO MALE WOLFFIAN DUCT INCLUDES SURROUNDING  
region cDNA, RIKEN full-length enriched library, clone:6720483C06  
product: plasma glutamate carboxypeptidase, full insert sequence.  
AK032972  
VERSION  
AK032972.1 GI:26328732  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
9879253  
MEDLINE  
PUBMED  
10349636  
AUTHORS  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to

JOURNAL  
MEDLINE  
PUBMED  
20499374  
11042159  
AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
20530913  
11076861  
AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
21085660  
11217851  
AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
21085660  
11217851  
AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
21085660  
11217851  
AUTHORS

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
21085660  
11217851  
AUTHORS

COMMENT

prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,  
Kono, H., Akiyama, J., Nishi, K., Kikunishi, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, S.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Caaveant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nakai, C., Nishikawa, G.,  
Quackenbush, J., Schriml, L. M., Steinhilber, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D. A., Kamuya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, D., Mombauts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Tayo-Oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S.  
and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 1739)  
Adachi, J., Aizawa, K., Akimura, T., Aikawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirose, T.,  
Hori, F., Imocant, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki-Ahira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki. The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan [E-mail: genome-res@gsr.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216]  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in Riken  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

FEATURES  
source  
1..1739  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:6720483C06"  
/db\_xref="taxon:10090"  
/clone="6720483C06"  
/sex="male"  
/tissue\_type="olfactory duct includes surrounding region"  
/clone\_id="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="12 days embryo"  
81..1493  
/note="unnamed protein product; plasma glutamate  
carboxypeptidase (MGI:1889205, GI:1889205, GI:1889205, evidence:  
BLASTN, 98%, match=943)  
putative"  
/protein\_id="BAC28105.1"  
/cdon\_start=1  
/trnslatlon="MRSLEFLVHLALGSKAVKNGVORTPREIKEELANYEDV  
AKAIIIMAVYKXONYSERILGLVDVTPGRLSGKRLKALIDIMYONQOQDELNVH  
LEQVRIPIHMERGESAVMEPRIRKHAALIGSSIGTPPGIATAVLVASPELGR  
ASERAKIIVNPNYGVETKTVYRQVGAKEAAGVAVSLQSVASPSISPHGIG  
KYDDGPKPIPTACTITVEDAEMSRMSRKNKVIYIH.EMKATYPTDPSPTVAEIGS  
MYEEVLVLSNDVQGLDDGAGATSEALSLVLDLRLKRLTVNLVLP  
EEOGIAISQYVEHLKNDISQYSLVMDADSGTLPGLQFGSDGALIMKENVNLQ  
LNTAKVFSNGEGDINFWIOAGVPSALRDLDTLYKFFHSHSDTVMQKQNVAA  
AVWAVVAVVADMDENLPRS"  
1722..1727  
/note="putative"  
polya\_site  
1739  
/note="putative"

CDS  
81..1493  
/note="unnamed protein product; plasma glutamate  
carboxypeptidase (MGI:1889205, GI:1889205, GI:1889205, evidence:  
BLASTN, 98%, match=943)  
putative"  
/protein\_id="BAC28105.1"  
/cdon\_start=1  
/trnslatlon="MRSLEFLVHLALGSKAVKNGVORTPREIKEELANYEDV  
AKAIIIMAVYKXONYSERILGLVDVTPGRLSGKRLKALIDIMYONQOQDELNVH  
LEQVRIPIHMERGESAVMEPRIRKHAALIGSSIGTPPGIATAVLVASPELGR  
ASERAKIIVNPNYGVETKTVYRQVGAKEAAGVAVSLQSVASPSISPHGIG  
KYDDGPKPIPTACTITVEDAEMSRMSRKNKVIYIH.EMKATYPTDPSPTVAEIGS  
MYEEVLVLSNDVQGLDDGAGATSEALSLVLDLRLKRLTVNLVLP  
EEOGIAISQYVEHLKNDISQYSLVMDADSGTLPGLQFGSDGALIMKENVNLQ  
LNTAKVFSNGEGDINFWIOAGVPSALRDLDTLYKFFHSHSDTVMQKQNVAA  
AVWAVVAVVADMDENLPRS"  
1722..1727  
/note="putative"

BASE COUNT 482 a 369 c 423 g 465 t  
ORIGIN

Query Match 80.8%; Score 1086.4; DB 11; Length 1739;  
Best Local Similarity 88.0%; Pred. No. 0;  
Matches 1183; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

1 AAGATGGCATCTCTAAGAGACCTTTGAAGAATAAAGAAATAGCAGCTGGA 60  
147 AAGATGGCTTTCTCGACCAATTTGGAATAAAGAAATAGCAGCTGGA 206  
61 GATGTGCTAAGCAATCATCACTAGCTGTTATGTTAAGCCGAGACAGATCCTAT 120  
207 GATGTGCTAAGCAATCATCACTAGCTGTTATGTTAAGCCGAGACAGATCCTAT 266  
121 GAGCGATGGCACTTCTGCTTGAATCTGTTGAACCCAGACTGAGTGGCTCAAGAACCTTA 180  
267 GAGCGTGGGAGCTTCTGCTTGAATCTGTTGAACCCAGACTGAGTGGCTCAAGAACCTTA 326  
181 GAAAGAGCATCAATTAATGTTCAAAACCTGAGAGAAATAGGAGTGGAGTGGAC 240  
327 GAGAAACCTTCAATCAATGTTCAAAACCTGAGAGAAATAGGAGTGGAGTGGAC 386  
241 CTGAGAGCATGAGATATCCCACTGAGAGAGGAGAGAAATAGCTGATGCTGGAG 300  
387 CTGAGAGCATGAGATATCCCACTGAGAGAGGAGAGAAATAGCTGATGCTGGAG 446  
301 CCAAGATTAATTAAGATAGCCTCTGAGTGGAGAGAGATGAGTGGAGTGGAGTGGAG 360  
447 CTTGATTAATTAAGATAGCCTCTGAGTGGAGAGAGATGAGTGGAGTGGAGTGGAG 506  
361 GGAATTAACAGAGATGTTGTTATTAACCACTTACATCACTACAGAGAGAGTGG 420  
507 GGAATTAACAGAGATGTTGTTATTAACCACTTACATCACTACAGAGAGAGTGG 566  
421 GGAATTAACAGAGATGTTGTTATTAACCACTTACATCACTACAGAGAGAGTGG 480

Db	567	GAAGCAAGAGGAGATCATTTGTTATTAACACAGCCTTACATGCTATGAGAAAGACTGTG	626
Qy	481	CAATACCAAGCGAGCGGCGGTGAGACTGCGCAAGGTGGGGGCTTTGGCATCTCTCAATT	540
Db	627	CAGATCCGGGTGCAGAGAGCTGTGAGACTGCGCAAGGTGGGGGCTGTGGCAATCCCTCAATC	686
Qy	541	CGATCCGGGTGCTCTCTCCATCTACAGTCTCTCACAGAGGATTCAGAGAAATACAGAGAT	600
Db	687	CAATCAGTAGCTCTCTTTCCATCTACAGTCTCTCACAGAGGATTCAGAGAAATATCAAGAT	746
Qy	601	GCGGTGCCCAAAATTCACAGCCTGTATTAACGCTGAGAGATGCGAAATAGTCAAGA	660
Db	747	GGTGGCCCAAGATTCACAGCCTGTATTCACAGTGAAGATGCGAAATAGTCTCGA	806
Qy	661	ATGGTTCTTCATGGGATCAAAATTTCTACATGCTGAAGATGGGGGCGCAAGCCATACCA	720
Db	807	ATGGTTCTTCGAGGAAACAAATTTCTACATGCTGAAGATGGGGGCGCAAGCCATATCA	866
Qy	721	GATACCTGATTCCTTCACACATGCTGACAGATCACTGGGAGCAATATTCAGAACAGATT	780
Db	867	GATACCTGATTCCTTCACATGCTGACAGATCACTGGGAGCAATATTCAGAACAGATT	926
Qy	781	GTACTGTGCTGAGCATCTGAGCAGCTGGAGTGTGGGCGAGGTGCGATGATGATGATGATG	840
Db	927	GTCCTGTGCTGAGCATCTGAGCAGCTGGAGTGTGGGCGAGGTGCGATGATGATGATGATG	986
Qy	841	GCTGAGCCTTTATATCATGGAAGCACTCTCTATTAAGATCTTGGGTGGTGTCTCA	900
Db	987	GCTGAGCCTTTATATCATGGAAGCACTCTCTATTAAGATCTTGGGTGGTGTCTCA	1046
Qy	901	AAGAGACTCTGCGGCTGCTCTGAGCTGAGCAGAGAAACAAGGTGAGTGGTCTTC	960
Db	1047	AAGAGACTCTGCGGCTGCTGCTGAGCTGAGCAGAGAAACAAGGTGAGTGGTCTTC	1106
Qy	961	CAGTATATCAGTTACACAGGTAATATTTCCACTACAGTCTGTGATGAGTGTGATC	1020
Db	1107	CAGTATATGAGCTACATTAAGCAATATTTCCAGTACAGTCTGTGATGAGTGTGATC	1166
Qy	1021	GCAGGAACCTTTTCCACTGGGTGCAATTCATCTGGCAGTGAAGAAAGGCCAGGCCATC	1080
Db	1167	TCAGGAACCTTTTCCACTGGGTGCAATTCATCTGGCAGTGAAGAAAGGCCAGGCCATC	1226
Qy	1081	ATGAGGAGGTTATGAGCTGCTGAGCCCTCAATATCACTGCTGAGCCATGGA	1140
Db	1227	ATGAGGAGGTTATGAGCTGCTGAGCCCTCAATATCACTGCTGAGCCATGGA	1286
Qy	1141	GAAGGAGACATCACTTTTGGATCCAAAGCTGAGTGTGAGCCAGTCTACTTAT	1200
Db	1287	GAAGGAGACATCACTTTTGGATCCAAAGCTGAGTGTGAGCCAGTCTACTTAT	1346
Qy	1201	GACTTATCAAGTATTTCTTCCATCACTGCGCAGAGACACATGATGCTCATGAT	1260
Db	1347	GACTTATCAAGTATTTCTTCCATCACTGCGCAGAGACACATGATGCTCATGAT	1406
Qy	1261	CCAAAGCAGATGATGCTGCTGCTGTTGGGCTGTGTTCTTATGTTGTCAGAC	1320
Db	1407	CCAAAGCAGATGATGCTGCTGCTGTTGGGCTGTGTTCTTATGTTGTCAGAC	1466
Qy	1321	ATGGAAGAAATCTGCTAGGTGC	1344
Db	1467	ATGGAAGAAATCTGCTAGGTGC	1490

RESULT 3  
AK075686  
LOCUS  
DEFINITION  
MUS musculus 18-day embryo whole body cDNA, RIKEN full-length  
cDNA library, clone:119003P12 product:plasma glutamate  
carboxypeptidase, full insert sequence.

ACCESSION  
AK075686  
VERSION  
AK075686.1  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)

ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionath; Muridae; Murinae; Mus.
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitamura, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Stadli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, D., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mondaets, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wysshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohetsuki, S. and Hayashizaki, Y.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409 (6821), 685-690 (2001)
MEDLINE	21085660
PUBMED	11217851
REFERENCE	5
AUTHORS	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 1805)
AUTHORS	Aizawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hizumoto, K., Hirakawa, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasekawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Konda, M., Koyama, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shitaki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Teijima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

**TITLE** Direct Submission  
**JOURNAL** Submitted (16-APR-2002) Yoshihide Hayashizaki, the Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
**COMMENT** cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.go.jp/  
 URL: http://fantom.gsc.riken.go.jp/  
**FEATURES** Location/Qualifiers  
**SOURCE** 1. 1805  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="PANTOM,DB:1190003P12"  
 /db\_xref="MGI:1909852"  
 /db\_xref="taxon:10090"  
 /clone="1190003P12"  
 /tissue\_type="whole body"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="18-day embryo"  
 147..1559  
 /note="unnamed protein product; plasma glutamate carboxypeptidase (MGP|MGI:1889205, GB|NM\_018755, evidence: BLASTN, 99%, match=943)  
 putative"  
 /codon\_start=1  
 /protein\_id="BAC35891.1"  
 /db\_xref="MGI:1923585"  
 /db\_xref="GI:26344481"  
 /translation="MRSLFLFIVHLALSGKAVFKNGVSVORTFEIKIEIANVEDV AKAINLAVGKYONRSYERGLIVDPVLPGLSGSKLAKAIOIMYONLOODGLENVH LEQVRIPIHMERGESAVMLERHKNAILIGSSISGPRPGITAEVIVVASFDELQKR ASERAKIITYNDYIEKTVQYRVGSAVEAKVAVASLIQSVASFISYPTGIGIQ KYDGVKPIACTIVEDAEWMSMARSKNIVILHMGAKTYDPTDSFNTVAITGS MYPEEVALVSGHDSMDVQGLDDGGAGKIFSMELSLVXDIGARPRTILVLTIGS EOGGIGASOYVELHKANISKYSLVMEADSGTFPLGQFOTGSDKARAIMKEVNMILQP LNTKVFNSGEGTDINFMIQAGVPSASLRDLVYKFFEFHSHSDTMTVMKQNNVAA AVNAVVAIVVADMDMLPRIS"  
 1788..1793  
 /note="putative"  
 1805  
 /note="putative"  
**BASE COUNT** 486 a 397 c 445 g 476 t 1 others  
**ORIGIN**  
 Query Match 80.7%; Score 1084.8; DB 11; Length 1805;  
 Best Local Similarity 87.9%; Pred. No. 0;  
 Matches 1182; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

241 CTGAGCCAGTGAATATCCCACTGGAGAGGGGAGAAATACGTGTGATGCTGGAG 300  
 Db 453 CTGAGAGAGGTCAAGATATCCCACTGGAGAGGGGAGAAATCTGAGTATGCTTGG 512  
 Qy 301 CCAAGATTCATTAAGATGACATCTGGGTCTTGGACGACGATTTGGATCTCTCCAGAA 360  
 Db 513 CCTCGAATTCACAAAGAGGTATCTAGGCTTGGACGACGATTTGGATCTCTCCAGAA 572  
 Qy 361 GGCATTCACAGAAAGTGTGTGTGTGACCTCTTGTGATGAATCTGACAGAGGGCTCA 420  
 Db 573 GGCATTCACAGAAAGTGTGTGTGTGACCTCTTGTGATGAATCTTCAAGAGAGATCA 632  
 Qy 421 GAAGCAAGAGAGAAATTTGTTTATTAACCACTTACATCACTCAAGAGCGGTG 480  
 Db 633 GAAGCAAGAGAGAAATTTGTTTATTAACCACTTACATCACTGAGTATGAAGACTGTG 692  
 Qy 481 CAATACCGAAACGAGAGGGGCGGTGGAAGCTGCAAGGTGGGGCTTTGGATCTCTATT 540  
 Db 693 CAGTACCGGGTGCAGGAGCTGTGGAAGCTGCCAAGGTGGAGCTGTGGATCTCTATC 752  
 Qy 541 CGATCCGGTGCCTCTCTCATCTACAGTCTCACAAGGTATTCAGAAATCCAGAGT 600  
 Db 753 CAATCAGTACCTCTCTCTCATCTACAGTCTCACAAGGTATTCAGAAATCCAGAGT 812  
 Qy 601 GGCCTGCCAAATTCACAGAGCTGTATTACGGTGGAGATGCAGAAATGATGTCAAGA 660  
 Db 813 GGTGTGCCAAGATTCACAGAGCTGTATTACGGTGGAGATGCAGAAATGATGTCTCGA 872  
 Qy 661 ATGGCTTTCATGGGATCAAAATTTGTCATTCAGCTAAAGATGGGGCAAGACTTACCA 720  
 Db 873 ATGGCTTTCATGGGATCAAAATTTGTCATTCAGCTAAAGATGGGGCAAGACTTACCA 932  
 Qy 721 GATATGATTCCTTCAACACTGTAGAGATCACTGGGAGCAAAATTCAGAAACAGTGT 780  
 Db 933 GATATGATTCCTTCAACACTGTAGAGATCACTGGGAGCAAAATTCAGAAACAGTGT 992  
 Qy 781 GTAATGATTCCTTCAACACTGTAGAGATCACTGGGAGCAAAATTCAGAAACAGTGT 840  
 Db 993 GTCTGTGATGATGACATTTGGACAGTGGGATGTTGGGACGATGATGATGATGATG 1052  
 Qy 841 GGTGAGCTCTTATATCATGAGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCA 900  
 Db 1053 GGTGAGCTCTTATATCATGAGGAAGCACTCTCACTTATTAAGATCTTGGGCTGCTCA 1112  
 Qy 901 AAGAGACTGTGGGCTGTGTCTGTGACTGCAAGAAAGCAAGGAAATTTGCTCTCC 960  
 Db 1113 AAGAGACTGTGGGCTGTGTCTGTGACTGCAAGAAAGCAAGGAAATTTGCTCTCC 1172  
 Qy 961 CAGTATTCATGATTAACAAGATTAATTTCCACTAGCTGCTGGATGAGAGCTGAC 1020  
 Db 1173 CAGTATTCATGATTAACAAGATTAATTTCCACTAGCTGCTGGATGAGAGCTGAC 1232  
 Qy 1021 GCAGGAACCTTCTTACCCACTGGGCTGCAATTCATCTGCAAGTGAAGGCAAGGCCATC 1080  
 Db 1233 TCAGGAACCTTCTTACCCACTGGGCTGCAATTCATCTGCAAGTGAAGGCAAGGCCATC 1292  
 Qy 1081 ATGAGAGAGTATGAGCTGCTGCTGCAAGCTTCAATATCACTAGCTGCTGAGCATGGA 1140  
 Db 1293 ATGAGAGAGTATGAGCTGCTGCTGCAAGCTTCAATATCACTAGCTGCTGAGCATGGA 1352  
 Qy 1141 GAAGGAGACAGATCAATCTTGTGATCCAGAGCTGAGTGTGCTGAGGACGATCTACTGAT 1200  
 Db 1353 GAAGGAGACAGATCAATCTTGTGATCCAGAGCTGAGTGTGCTGAGGACGATCTACTGAT 1412  
 Qy 1201 GACTTATACAAAGTATTTCTTCTTCATCACTCCACGAGACACATGACTGTCTATGAT 1260  
 Db 1413 GACTTATACAAAGTATTTCTTCTTCATCACTCCACGAGACACATGACTGTCTATGAT 1472  
 Qy 1261 CCAAGAGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAC 1320  
 Db 1473 CCAAGAGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAC 1532  
 Qy 1321 ATGGAAGAAATGCTGCTAGGTCC 1344



DB 1533 ATGATGAATGCTGCCAGCTCC 1556

|||||

RESULT 4  
BC012019  
LOCUS  
DEFINITION Homo sapiens, similar to plasma glutamate carboxypeptidase, clone  
IMAGE:4455631, mRNA.  
ACCESSION EC012019.1 GI:15080558  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Strausberg, R.  
TITLE 1 (bases 1 to 1024)  
JOURNAL Submitted (30-JUL-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
REMARK  
COMMENT Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnarone, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Louised, H.,  
Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>  
Series: IRAX Plate: 28 Row: b Column: 11  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 5174626  
This clone has the following problem: retained intron.

FEATURES  
source  
1..1024  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4455631"  
/cissue\_type="Kidney, hypernephroma"  
/clone\_id="NIH\_MGC\_89"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
BASE COUNT 317 a 210 c 256 g 241 t  
ORIGIN

Query Match 66.0%; Score 886.8; DB 11; Length 1024;  
Best local Similarity 99.8%; Pred. No. 76-255;  
Matches 888; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AAGATGCAATCTCAAGAGGAGCTTTGAAGAAATGAAGAAATGAGCAAGCTGTGGA 60  
DB 108 AAGATGCAATCTCAAGAGGAGCTTTGAAGAAATGAAGAAATGAGCAAGCTGTGGA 167  
QY 61 GATGTTGCTAAAGCAATCACTAGCTGTTTATGTTAAAGCCAGAGCATCTAT 120  
DB 168 GATGTTGCTAAAGCAATCACTAGCTGTTTATGTTAAAGCCAGAGCATCTAT 227  
QY 121 GAGGATTTGCACTTCTGTTGATACCTGTTGAGCCAGAGCTGAGTGTCCAGAACTTA 180  
DB 228 GAGGATTTGCACTTCTGTTGATACCTGTTGAGCCAGAGCTGAGTGTCCAGAACTTA 287

QY 181 GAAAAGCATCCAAATTAATGATCAAAACCTGAGCAAGATGGCTGAGAAAGTTTAC 240  
DB 288 GAAAAGCATCCAAATTAATGATCAAAACCTGAGCAAGATGGCTGAGAAAGTTTAC 347  
QY 241 CTGAGCCAGTGAATACCCCACTGGAGAGGGAGAGAAATCAGTGTGATGCTGAG 300  
DB 348 CTGAGCCAGTGAATACCCCACTGGAGAGGGAGAGAAATCAGTGTGATGCTGAG 407  
QY 301 CCAAGATTCATTAAGATAGCATTCCTGGTCTTGGCAGCAGCATTTGGATCTCCCGAA 360  
DB 408 CCAAGATTCATTAAGATAGCATTCCTGGTCTTGGCAGCAGCATTTGGATCTCCCGAA 467  
QY 361 GGCATTACAGAGAGTCTGTTGTTGATGATGATGATGATGATGATGATGATGATGAT 420  
DB 468 GGCATTACAGAGAGTCTGTTGTTGATGATGATGATGATGATGATGATGATGATGAT 527  
QY 421 GAACCAAGAGAGAGTCTGTTGTTGATGATGATGATGATGATGATGATGATGATGAT 480  
DB 528 GAACCAAGAGAGAGTCTGTTGTTGATGATGATGATGATGATGATGATGATGATGAT 587  
QY 481 CAATACCGAAGCAGAGGG 540  
DB 588 CAATACCGAAGCAGAGGG 647  
QY 541 CGATCCGTGGCTCTCTCTCATCTACATCTACATCTACATCTACATCTACATCTACAT 600  
DB 648 CGATCCGTGGCTCTCTCTCATCTACATCTACATCTACATCTACATCTACATCTACAT 707  
QY 601 GGCCTGCTCAAGATTCGAACAGCTGTATTAAGCTGTAAGATGAGATGATGATGATGAT 660  
DB 708 GGCCTGCTCAAGATTCGAACAGCTGTATTAAGCTGTAAGATGAGATGATGATGATGAT 767  
QY 661 ATGCTCTCATGAGATCAAAATTTGATGATGATGATGATGATGATGATGATGATGAT 720  
DB 768 ATGCTCTCATGAGATCAAAATTTGATGATGATGATGATGATGATGATGATGATGAT 827  
QY 721 GATCTGATTCCTTCAACACTGTAGCAGATCACTGGAGCAAAATTTCCAGAACAGTT 780  
DB 828 GATCTGATTCCTTCAACACTGTAGCAGATCACTGGAGCAAAATTTCCAGAACAGTT 887  
QY 781 GATCTGATTCCTTCAACACTGTAGCAGATCACTGGAGCAAAATTTCCAGAACAGTT 840  
DB 888 GATCTGATTCCTTCAACACTGTAGCAGATCACTGGAGCAAAATTTCCAGAACAGTT 947  
QY 841 GGTGAGCTTTATATGATGAGAAAGCACTCTCACTTTTAAAGATCTTG 890  
DB 948 GGTGAGCTTTATATGATGAGAAAGCACTCTCACTTTTAAAGATCTTG 997

RESULT 5  
BX439467  
LOCUS  
DEFINITION Homo sapiens PLACENTA Homo sapiens cDNA clone CSDB010Y006  
5-PRIME, mRNA sequence.  
ACCESSION BX439467  
VERSION BX439467.1 GI:30777746  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), web: [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4663.f For  
more information about this cluster, see  
<http://www.genoscope.cns.fr/>



cg1-bin/cluster.cgi?seq=CS0DE010DE3QPlc1cluster=4663.f. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0DE010DE3QPl.

## FEATURES

SOURCE

Location/Qualifiers  
 1. 1038

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE010Y106"  
 /tissue\_type="PLACENTA"  
 /clone\_lib="Homo sapiens PLACENTA"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with Not I and cloned into  
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."  
 BASE COUNT 307 a 233 c 271 g 239 t  
 ORIGIN

Query Match 62.0%; Score 833.8; DB 13; Length 1050;  
 Best Local Similarity 99.8%; Pred. No. 6.2e-239;  
 Matches 835; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AAGAATGCGATCTCTAAGAGACCTTTGAAGAAATTAAGAAGAAATAGCCAGCTGTGA 60  
 214 AAGAATGCGATCTCTAAGAGACCTTTGAAGAAATTAAGAAGAAATAGCCAGCTGTGA 273  
 61 GATGTTGCTAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAGCAACAGATCTAT 120  
 274 GATGTTGCTAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAGCAACAGATCTAT 333  
 121 GAGGATTTGGAACCTCTGTTGATATCTGTGAGACCCAGACCTGAGTGGCTCCAGAACCTA 180  
 334 GAGGATTTGGAACCTCTGTTGATATCTGTGAGACCCAGACCTGAGTGGCTCCAGAACCTA 393  
 181 GAAAAAGCCATCCAAATTTATGTATGCCAAACCTGACAGCAAGATGGCTGAGAAAGTTTAC 240  
 394 GAAAAAGCCATCCAAATTTATGTATGCCAAACCTGACAGCAAGATGGCTGAGAAAGTTTAC 453  
 241 CTGAGCCAGTGAAGATATCCCACTGGAGAGAGGGAGAAATCACTGCTGTGATGCTGAG 300  
 454 CTGAGCCAGTGAAGATATCCCACTGGAGAGAGGGAGAAATCACTGCTGTGATGCTGAG 513  
 301 CCAAGAAATTCATTAAGATAGCACTCTGCTGCTGAGCAACAGATGGAGATCTCTCCAGAA 360  
 514 CCAAGAAATTCATTAAGATAGCACTCTGCTGCTGAGCAACAGATGGAGATCTCTCCAGAA 573  
 361 GGCATTACAGCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 574 GGCATTACAGCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 633  
 421 GAAGCAAGAGGAGATGTTGTTTATTAACCAACCTTACATCACTACTCAAGACGGTGA 480  
 634 GAAGCAAGAGGAGATGTTGTTTATTAACCAACCTTACATCACTACTCAAGACGGTGA 693  
 481 CAATTCAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
 694 CAATTCAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 753  
 541 CGATCCGAGGCTCTCTCTCATCTACAGTCTCTACAGGATTCAGGAATATACAGGAT 600  
 754 CGATCCGAGGCTCTCTCTCATCTACAGTCTCTACAGGATTCAGGAATATACAGGAT 813  
 601 GGCCTGCCAAATTCACAGCTGTTATTAAGTGAAGATGAGAAATGATGCAAGA 660  
 814 GGCCTGCCAAATTCACAGCTGTTATTAAGTGAAGATGAGAAATGATGCAAGA 873  
 661 ATGGCTTCATAGGATCAAAATTTGATCACTGCTTAAGATGGGGGCAAGACCTAACCA 720  
 874 ATGGCTTCATAGGATCAAAATTTGATCACTGCTTAAGATGGGGGCAAGACCTAACCA 933  
 721 GATACCTGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATTCAGAACAGTT 780

Db 934 GATACCTGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATTCAGAACAGTT 993  
 Qy 781 GATACCTGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATTCAGAACAGTT 837  
 Db 994 GATACCTGATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATTCAGAACAGTT 1050

RESULT 6  
 LOCUS BX355940 1038 bp mRNA linear EST 05-MAY-2003  
 DEFINITION BX355940 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA  
 clone CS0D1006Y115 5-PRIME, mRNA sequence.  
 ACCESSION BX355940  
 VERSION BX355940.1 GI:30384019  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to this cluster, see  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cg1-bin/cluster.cgi?seq=CS0D1006AB08QPlc1cluster=4663.f. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0D1006AB08QPl.

## FEATURES

SOURCE

Location/Qualifiers  
 1. 1038

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1006Y115"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoRV  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 304 a 237 c 262 g 233 t 2 others  
 ORIGIN

Query Match 57.4%; Score 771.8; DB 13; Length 1038;  
 Best Local Similarity 99.6%; Pred. No. 2.8e-220;  
 Matches 784; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

1 AAGAATGCGATCTCTAAGAGACCTTTGAAGAAATTAAGAAGAAATAGCCAGCTGTGA 60  
 252 AAGAATGCGATCTCTAAGAGACCTTTGAAGAAATTAAGAAGAAATAGCCAGCTGTGA 311  
 61 GATGTTGCTAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAGCAACAGATCTAT 120  
 312 GATGTTGCTAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAGCAACAGATCTAT 371  
 121 GAGGATTTGGAACCTCTGTTGATATCTGTGAGACCCAGACCTGAGTGGCTCCAGAACCTA 180  
 372 GAGGATTTGGAACCTCTGTTGATATCTGTGAGACCCAGACCTGAGTGGCTCCAGAACCTA 431  
 181 GAAAAAGCCATCCAAATTTATGTATGCCAAACCTGACAGCAAGATGGCTGAGAAAGTTTAC 240  
 432 GAAAAAGCCATCCAAATTTATGTATGCCAAACCTGACAGCAAGATGGCTGAGAAAGTTTAC 491  
 241 CTGAGCCAGTGAAGATATCCCACTGGAGAGAGGGAGAAATCACTGCTGTGATGCTGAG 300

Db	492	CTGAGCCAGTGTGAAATACCCACCTGGGAGAGGGAGAAAGATCAAGCTGTGATGCTGGAG	551
QY	301	CCAGAATATTCATTAAGATAGCCATCTCGGCTCTTGGCAGCAGCATTTGGAGCTCTCCAGAA	360
Db	552	CCAAATAATTCATTAAGATAGCCATCTCGGCTCTTGGCAGCAGCATTTGGAGCTCTCCAGAA	611
QY	361	GGCATTTACGACGAAAGTTCTGGTGGTGAACCTTTTCGATGAACCTGACGAAAGGGCTTCA	420
Db	612	GGCATTTACGACGAAAGTTCTGGTGGTGAACCTTTTCGATGAACCTGACGAAAGGGCTTCA	671
QY	421	GAAGCAAGGGGAGAAATTTGTTTATAACAACCTTACATCACTCACTACAGACGGTG	480
Db	672	GAAGCAAGGGGAGAAATTTGTTTATAACAACCTTACATCACTCACTACAGACGGTG	731
QY	481	CAATACCGAAACGACAGGGGGCGGTGAAAGCTGCGAAAGTGGGGGCTTTGGCATCTTCAT	540
Db	732	CAATACCGAAACGACAGGGGGCGGTGAAAGCTGCGAAAGTGGGGGCTTTGGCATCTTCAT	791
QY	541	CGATCCGTGGCTCTCTTCTCCATCTACAGTCTCTCACACAGGATTCAGGAATACAGAT	600
Db	792	CGATCCGTGGCTCTCTTCTCCATCTACAGTCTCTCACACAGGATTCAGGAATACAGAT	851
QY	601	GGCGTGCCCAAAATTCCAACAGCTGTATTACGGTGAAGATGCAAGAAATGATGTCAAGA	660
Db	852	GGCGTGCCCAAAATTCCAACAGCTGTATTACGGTGAAGATGCAAGAAATGATGTCAAGA	911
QY	661	ATGGCTTCTCATGGGATCAAAATTTGTCAATTCAGTAAAGAT-GGGGGCAAAAGACTTACC	719
Db	912	ATGGCTTCTCATGGGATCAAAATTTGTCAATTCAGTAAAGATGGGGGCAAAAGACTTACC	971
QY	720	AGATATCTATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCGAACAGT	779
Db	972	AGATATCTATTCCTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCGAACAGT	1031
QY	780	TGTACTG 786	
Db	1032	TGTACTG 1038	

[illegible]

```
FEATURES
source
location/Qualifiers
1. .866
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6185233"
/sex="male"
```

```

/tissue_type="dorsal root ganglia"
/deg_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_id="lupski_dorsal_root_ganglion"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
directionally cloned using the following adaptors:
5'-TCGACCACGCGCTCCG-3' and
5'-GACTAGTCTTAGATCGCGAGCGGCCCTT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
BASE COUNT      257 a      182 c      233 g      191 t      3 others
ORIGIN
Query Match      55.1%; Score 740.6; DB 13; Length 866;
Best Local Similarity 99.1%; Pred. No. 6.3e-211;
Matches 765; Conservative 0; Mismatches 5; Indels 2; Gaps 2

```

OY	1	TAGATGSCATCTCTTAAGAGACCTT--TGAAGAAATATAAAGAAATAGCAGCTGG	59
Db	9	AAGATGSCATCTCTTAAGAGACCTTTTGAAGAAATATAAAGAAATAGCAGCTGG	68
OY	60	AGATGTTGCTTAAGCAATCATCAACCTTAGCTGTTTATGTATAAGCCCAAGACATCTTA	119
Db	69	AGATGTTGCTTAAGCAATCATCAACCTTAGCTGTTTATGTATAAGCCCAAGACATCTTA	128
OY	120	TGAGCGATTGGCACTTCTGGTTGATACTGTTGACCCAGACTGAGTGGCTCCAMGAACTT	179
Db	129	TGAGCGATTGGCACTTCTGGTTGATAGTGTGAGCCAGACTGAGTGGCTCCAMGAACTT	188
OY	180	AGAAAAAGCCATCCAAATTTATGTACCAAAACCTGACGAAGATGGGCTGAGAAAGTTCA	239
Db	189	AGAAAAAGCCATCCAAATTTATGTACCAAAACCTGACGAAGATGGGCTGAGAAAGTTCA	248
OY	240	CCTGAGCCAGTGAATAATCCCACTGGGAGAGGGAGAAAGATCAGCTGTGATGCTGGA	299
Db	249	CCTGAGCCAGTGAATAATCCCACTGGGAGAGGGAGAAAGATCAGCTGTGATGCTGGA	308
OY	300	GCCAAAGATTATTAAGTATGCCATCCTGGGCTTTGGCAGCAGCATTTGGGACTCTCCAGA	359
Db	309	GCCAAAGATTATTAAGTATGCCATCCTGGGCTTTGGCAGCAGCATTTGGGACTCTCCAGA	368
OY	360	AGGCAATTACAGCAGAGTTCGTGAGTGACCTCTTTCGATGAATGACGACAGAAAGGAGCTC	419
Db	369	AGGCAATTACAGCAGAGTTCGTGAGTGAGCTCTTTCGATGAATGACGACAGAAAGGAGCTC	428
OY	420	AGAAGCAAGAGGGAAGATTGTTGTTTATTAACCAACTTACATCACTCACTCAAGAGCGGT	479
Db	429	AGAAGCAAGAGGGAAGATTGTTGTTTATTAACCAACTTACATCACTCACTCAAGAGCGGT	488
OY	480	GCAATACCGAACGCAAGGGGGCGGTGAAAGCTGCCAAGTGGGGGCTTTGGCATCTTCAT	539
Db	489	GCAATACCGAACGCAAGGGGGCGGTGAAAGCTGCCAAGTGGGGGCTTTGGCATCTTCAT	548
OY	540	TGCAATCGGTGGCCCTTTCATCTCAAGTCCACACAGGTATTTACAGAAATACCCAGA	599
Db	549	TGCAATCGGTGGCCCTTTCATCTCAAGTCCACACAGGTATTTACAGAAATACCCAGA	608
OY	600	TGGCGTCCCAAAATTTCCACAGCCTGTATTTACGGTGGAGATCAGAAATAGATGTCAAG	659
Db	609	TGGCGTCCCAAGATTTCCACAGCCTGTATTTACGGTGGAGATTCAGAAATAGATGTCAAG	668
OY	660	AATGGCTTCTCATGGGATCAAAATTTGCATTTCACTAAAGATGGGGGCAAAAGACTTACC	719
Db	669	AATGGCTTCTCATGGGATCAAAATTTGCATTTCACTAAAGATGGGGGCAAAAGACTTACC	728
OY	720	AGATACGATTCTTTCAACACT--GTACAGAGATCACCTGGGAGCAAAATATCTC	770
Db	729	AGATACGATTCTTTCAACACTGGTATACAGAGATCACCTGGGAGCAAAATATCTC	780

RESULT 8  
BG751497 902 bp mRNA linear EST 15-MAY-2001  
LOCUS 60273007F1 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:4873770 5',  
DEFINITION mRNA sequence.  
ACCESSION BG751497  
VERSION BG751497.1 GI:14062150  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 902)  
NIH-MGC <http://mgs.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-romail.nih.gov](mailto:cgaps-romail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/MLNL at:  
<http://image.llnl.gov>  
Plate: L1CM1752 row: 9 column: 19  
High quality sequence stop: 817.  
Location/Qualifiers  
1..902  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4873770"  
/tissue\_type="normal pigmented retinal epithelium"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH\_MGC\_43"  
/notes="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adapter: GGCAAGAG(G). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

BASE COUNT 246 a 200 c 251 g 204 t 1 others

ORIGIN

Query Match 55.1%; Score 740; DB 10; Length 902;  
Best Local Similarity 95.7%; Pred. No. 9.6e-211;  
Matches 847; Conservative 0; Mismatches 26; Indels 12; Gaps 8;

QY 259 CCCCACTGGAGAGAGAGAGAAATCACTGATGCTGAGCCAGAAATTCATTAAGATA 318  
DB 2 CCCCACTGGAGAGAGAGAGAGAAATCACTGATGCTGAGCCAGAAATTCATTAAGATA 61  
QY 319 GCCATCTGGGCTCTTGGCAGACGATGGGACTCTCTCAGAAAGCATTAACGCAAGATT 378  
DB 62 GCCATCTGGGCTCTTGGCAGACGATGGGACTCTCTCAGAAAGCATTAACGCAAGATT 121  
QY 379 CTGGTGGTGAACCTCTTTCGATGATGACGAGAGAGGGCTCAGAAAGCAAGAGGGAATT 438  
DB 122 CTGGTGGTGAACCTCTTTCGATGATGACGAGAGAGGGCTCAGAAAGCAAGAGGGAATT 180  
QY 439 GTTGTATTAAACAACCTTACATCACTCAAGAGAGGATGCAATACCAAGCAGAGGG 498  
DB 181 GTTGTATTAAACAACCTTACATCACTCAAGAGAGGATGCAATACCAAGCAGAGGG 240  
QY 499 GCGGTGAAGCTGCAAGAGTGGGGCTTGGCATCTCATTTGATCGTGGCTCTCTTC 558  
DB 241 GCGGTGAAGCTGCAAGAGTGGGGCTTGGCATCTCATTTGATCGTGGCTCTCTTC 300  
QY 559 TCATCTTAAGTCTCAACAGATATTCAAGAAATACAGATGGGCTGCCAATAATTCGA 618

DB 301 TCATCTTAAGTCTCAACAGATATTCAAGAAATACAGATGGGCTGCCAATAATTCGA 360  
QY 619 ACAGCTCTGATTTACGGTGAAGATGCAAAATGATGCAAGATAGCTTCTCATGGGATC 678  
DB 361 ACAGCTCTGATTTACGGTGAAGATGCAAAATGATGCAAGATAGCTTCTCATGGGATC 420  
QY 679 AAATGTCACTTACGATTAAGATGGGGGCAAGAGCTTACCAGATGATGATTCCTTCAAC 738  
DB 421 AAATGTCACTTACGATTAAGATGGGGGCAAGAGCTTACCAGATGATGATTCCTTCAAC 480  
QY 739 ACTGTACAGAGATCACTGTGAGAGCAATATCCAGAACAGGTTGACTGTGATGAGCAT 798  
DB 481 ACTGTACAGAGATCACTGTGAGAGCAATATCCAGAACAGGTTGACTGTGATGAGCAT 540  
QY 799 CTGACACGCTGGGATGTTGGGCGAGGGTCCATGG-ATGATGGCGGTGAGGCTTTATATC 857  
DB 541 CTGACACGCTGGGATGTTGGGCGAGGGTCCATGGCATGATGGCGGTGAGGCTTTATATC 600  
QY 858 ATGGAGAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTGCGGCT 917  
DB 601 ATGGAGAGCACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTGCGGCT 660  
QY 918 GTGTCTCTGACT-GCAGAAAGCAAGTGGAGTGGCTTCCAGTATTTATCATGTTAC 976  
DB 661 GTGTCTCTGACTGGAGAGAGCAAGTGGAGTGGCTTCCAGTATTTATCATGTTAC 720  
QY 977 ACAAGTAAATATTTCCACTA--CAGTCTGTGATGAGTTC-TGACGCGAGAACTTTCT 1033  
DB 721 ACAAGTAAATATTTCCACTAAGTCTGTGATGAGTTC-TGACGCGAGAACTTTCT 780  
QY 1034 T---ACCCACTGGGCTGCAATTTAC--TGGCAGTGAAGAGCCAGGCGCATCATGAGAG 1089  
DB 781 TTAACCACTGGGCTGCAATTTACCTTGGCAGTGAAGAGCCAGGCGCATCATGAGAG 840  
QY 1090 GTTATGAGCTGTGTCAGCCCTCAATATCACTAGGCTCTGAGC 1134  
DB 841 GTTATGAGCTGTGATGAAGCCC--TCATATCACTAGGCTCCGAAC 883

RESULT 9  
BX355939/c 996 bp mRNA linear EST 05-MAY-2003  
LOCUS BX355939 Homo sapiens P1ACENTRA.COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CSOD1006115 3-PRIME, mRNA sequence.  
ACCESSION BX355939.1 GI:30382027  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 996)  
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4663.f For  
more information about this cluster, see  
<http://www.genoscope.cns.fr/>  
cgi-bin/cluster.cgi?seq=CSOD1006A08NP1&cluster=4663.f. Contact :  
Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :  
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CSOD1006A08NP1.  
Location/Qualifiers  
1..996  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"



Qy 361 GGCAATTAAGAGCAAGATTCTGTGTGACCTCTTTCATGTAAGTGCAGAGAGGCGCTCA 420  
Db 496 GGCAATTAAGAGCAAGATTCTGTGTGACCTCTTTCATGTAAGTGCAGAGAGGCGCTCA 555  
Qy 421 GAAGCAAGAGGGAAGATTGTTTATTAACCACTTACATCACTACTCAAGAGCGGTG 480  
Db 556 GAAGCAAGAGGGAAGATTGTTTATTAACCACTTACATCACTACTCAAGAGCGGTG 615  
Qy 481 CAATACCGAAGCGAGGGGGCGGTGAGCTGCAAGGCGGGGGCTTTGGCATCTCTCAT 540  
Db 616 CAATACCGAAGCGAGGGGGCGGTGAGCTGCAAGGCGGGGGCTTTGGCATCTCTCAT 675  
Qy 541 CGATCGGTGGCTCTCTCTCATCTCAAGCTCTCAAGGATTTACAGAAATACAGAGAT 600  
Db 676 CGATCGGTGGCTCTCTCTCATCTCAAGCTCTCAAGGATTTACAGAAATACAGAGAT 735  
Qy 601 GGGGTGCCCCAAATTTCCAAAGCCTGTATTACGGTGAAGATGCAAAATATGTCAAGA 660  
Db 736 GGGGTGCCCCAAAGATTCAACAGCCTGTATTACGGTGAAGATGCAAAATATGTCAAGA 795  
Qy 661 ATGGCTTCTCATGGGATCAAAATTTGTCTCATGAGTAAAGAT-GGGGGCGAAGACCTAC 719  
Db 796 ATGGCTTCTCATGGGATCAAAATTTGTCTCATGAGTAAAGATGGGGGCGAAGACCTAC 855  
Qy 720 AGATACGATTCCTTCAACACT-GTAGCAGAGATCACTGG 759  
Db 856 AGATACGATTCCTTCAACACTGGTAGCAGAGATCACTGG 896

RESULT 11  
CB958693 739 bp mRNA linear EST 29-APR-2003  
LOCUS AGENCOURT 13778385 NIH MGC\_184 Homo sapiens cDNA clone  
DEFINITION IMAGE:30351744 5', mRNA sequence.  
ACCESSION CB958693  
VERSION CB958693.1 GI:30214809  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 739)  
NIH-MGC http://mgs.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-remail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovics  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: NDCM149 row: m column: 01  
High quality sequence stop: 572.  
Location/Qualifiers  
1. 739

FEATURES  
SOURCE  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30351744"  
/lab\_host="DH10B (TI phage-resistant)"  
/clone\_id="NIH\_MGC\_184"  
/note="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site 1:  
SfiI (ggccatcagcc); Site 2: SfiI (ggccggccggcc);  
Library is oligo-dr primed and directionally cloned. cDNA  
was prepared from a glandular pool of tissues from thyroid,  
parathyroid, adrenal, cortex and pineal gland. 5' and 3'  
adaptors were used in cloning as follows: 5' adaptor  
sequence: 5'-CACGGCCATTAATGGCC-3' and 3' adaptor sequence:  
5'-ATTTCAGAGCCGAGCGCCGACATG-dT(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.38

kb (range 0.60-3.5 kb). 15/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA). Note: this is a NIH-MGC Library."

BASE COUNT 192 a 171 c 195 g 181 t  
ORIGIN  
Query Match 53.5%; Score 719.2; DB 14; Length 739;  
Best Local Similarity 99.6%; Pred. No. 1.6e-204;  
Matches 721; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 516 GGTGGGGCTTTGGCATCTCTCATTCGATCCGTGGCTCTCTCTCACTACAGTCTCA 575  
Db 3 GGTGGGGCTTTGGCATCTCTCATTCGATCCGTGGCTCTCTCTCACTACAGTCTCA 62  
Qy 576 CACAGTATTCAGGAATACCAAGATGGCGTCCCAAAATTTCAACAGCTGTATTAGGT 635  
Db 63 CACAGTATTCAGGAATACCAAGATGGCGTCCCAAAATTTCAACAGCTGTATTAGGT 122  
Qy 636 GGAAGATGCAAGAAATGATGTCAAGATGGCTTCTCATGGGATCAAAATTTGTCATCAAGT 695  
Db 123 GGAAGATGCAAGAAATGATGTCAAGATGGCTTCTCATGGGATCAAAATTTGTCATCAAGT 182  
Qy 696 AAGATGGGGCAAAAGACTTACCAGATCTGATTTCTTCAACACTGTAGCAGAGATCAC 755  
Db 183 AAGATGGGGCAAAAGACTTACCAGATCTGATTTCTTCAACACTGTAGCAGAGATCAC 242  
Qy 756 TGGGAGCAATATTCAGAAAGGTTGTACTGTGTGATGATCTGACAGCTGGATGT 815  
Db 243 TGGGAGCAATATTCAGAAAGGTTGTACTGTGTGATGATCTGACAGCTGGATGT 302  
Qy 816 TGGGAGGAGGTCATGATGATGAGGAGGCTTTATATCAAGGAGAGCACTCTCACT 875  
Db 303 TGGGAGGAGGTCATGATGATGAGGAGGCTTTATATCAAGGAGAGCACTCTCACT 362  
Qy 876 TATTAAGATCTTGGGCTGGCTCCAAAGAGACTCTGGCTGGCTTGTGACTGACAG 935  
Db 363 TATTAAGATCTTGGGCTGGCTCCAAAGAGACTCTGGCTGGCTTGTGACTGACAG 422  
Qy 936 AGAACAAGTGTGAGTGTGGCTTCCAGATTTATTCATCAACAGGTAATATTTCCAA 995  
Db 423 AGAACAAGTGTGAGTGTGGCTTCCAGATTTATTCATCAACAGGTAATATTTCCAA 482  
Qy 996 CTACAGTCTGTGATGAGTGTGAGGAGAACTTCTTCCCACTGGCGTGCATTCAC 1055  
Db 483 CTACAGTCTGTGATGAGTGTGAGGAGAACTTCTTACCACTGGCGTGCATTCAC 542  
Qy 1056 TGGCAGTGAAGAGCCAGGCGCATCATGAGAGGATTATGAGCTGTGCAGCCCTCAA 1115  
Db 543 TGGCAGTGAAGAGCCAGGCGCATCATGAGAGGATTATGAGCTGTGCAGCCCTCAA 602  
Qy 1116 TATCACTCAGGTCTGTGAGCCATGAGAGAGGACAGACATCACTTTTGATTCAGCTGG 1175  
Db 603 TATCACTCAGGTCTGTGAGCCATGAGAGAGGACAGACATCACTTTTGATTCAGCTGG 662  
Qy 1176 AGTCCCTGAGCCAGCTCTACTGTGATCTTATACAGATTTCTTTCATCACTCCCA 1235  
Db 663 AGTCCCTGAGCCAGCTCTACTGTGATCTTATACAGATTTCTTTCATCACTCCCA 722  
Qy 1236 CGGA 1239  
Db 723 CGGA 726  
RESULT 12  
BX360507 1126 bp mRNA linear EST 05-MAY-2003  
LOCUS BX360507 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION BX360507 1126 bp mRNA linear EST 05-MAY-2003  
ACCESSION BX360507  
VERSION BX360507.1 GI:30374439  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

QY	541	CGATCCGGGCTCTCTCCATCTTACAGTCCCTACACAGGTATTTACAGAAATACAGGAT	600
Db	822	CGATCCGGGCTCTCTCTCCATCTTACAGTCCCTACACAGGTATTTACAGAAATACAGGAT	881
QY	601	GGCGTGGCCAAATATTCACACAGCCTGTATTACGT - G9AAGATGCAGAAATGATGTCAAG	659
Db	882	GGCGTGGCCAAAGATTTCACACAGCCTGTATTACGTGTGGGAAGATGCAGAAATGATGTCAAG	941
QY	660	AATGCTCTTCTCA - TGGGATCAAAATTTGTCATTTCAGCTTAAAGATGGGGCAAGACTTACC	718
Db	942	AATGCTCTTCTCATTTGGGATCAAAATTTGTCATTTCAGCTTAAAGATGGGGGGS - AARAMCTTACC	1000
QY	719	CAGATTCTGATTTCTTTCACACTGTGTACAGAGATCTGGAGAGAAATATTCACACAGG	778
Db	1001	CAGATTCTGATTTCTTTCACACTGTGTACAGAGATCTGGAGAGAAATATTCACACAGG	1058
QY	779	TTGTACTGTGT 788	
Db	1059	TACTGTACTGT 1068	
RESULT 13			
LOCUS	EXJ35995	919 bp	mrna linear EST 02-MAY-2003
DEFINITION	EXJ35995 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA		
ACCESSION	EXJ35995		
VERSION	EXJ35995.1	GI:30339459	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	Li W.B., Gruber C., Jesssee, J. and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
	Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr		
	Library was constructed by Life Technologies, a division of		
	Invitrogen. This sequence belongs to sequence cluster 4663.f. For		
	more information about this cluster, see		
	http://www.genoscope.cns.fr/		
	cgi-bin/cluster.cgi?seq=CS0D1022DE08Qp1&cluster=4663.f. Contact :		
	Feng Liang Email : fliang@life.techn.com URL : http://fulllength.invitrogen.com		
	http://fulllength.invitrogen.com/Invitrogen Corporation 1600		
	Paradey Avenue Genoscope sequence ID : CS0D1022DE08Qp1.		
FEATURES			
source	location/Qualifiers		
	1..919		
	/organism="Homo sapiens"		
	/mol_type="mrna"		
	/db_xref="taxon:9606"		
	/clone="CS0D1022YJ16"		
	/cruise_type="PLACENTA COT 25-NORMALIZED"		
	/cruise_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"		
	/note="1st strand cDNA was primed with a NotI-clio (dt)		
	primer. Five prime end enriched, double-strand cDNA was		
	digested with Not I and cloned into the Not I and EcoR V		
	sites of the pCMVSPORT 6 vector. Library was normalized."		
BASE COUNT	273 a 206 c 226 g 208 t		
ORIGIN			
Query Match	51.7%	Score 694.2;	DB 13; Length 919;
Best Local Similarity	98.3%;	Pred. No. 5.8e-197;	
Matches 707; Conservative	4;	Mismatches 7;	Indels 1; Gaps 1;
QY	1	AAGATGCGCATCTCTTAAGAGAGACTTTTGAAGAAATTAAGAAAGAAATAGCCAGCTGTGGA	60
Db	198	AAGATGCGCATCTCTTAAGAGAGACTTTTGAAGAAATTAAGAAAGAAATAGCCAGCTGTGGA	257

QY 61 GATGTCCTAAGCAATCATCACTAGCTGTTATAGTAAAGCCAGAAACAGATCTAT 120  
 DB 258 GATGTCCTAAGCAATCATCACTAGCTGTTATAGTAAAGCCAGAAACAGATCTAT 317  
 QY 121 GAGGATTTGGCACTTCTGTTGATACCTGTTGAGCCAGACTGATGGCTCCAAAGACTTA 180  
 DB 318 GAGGATTTGGCACTTCTGTTGATACCTGTTGAGCCAGACTGATGGCTCCAAAGACTTA 377  
 QY 181 GAAAAAGCCATCCAAATTTATGACCAAAACCTGAGAGAAATGGGCTGGAGAAAGTTTAC 240  
 DB 378 GAAAAAGCCATCCAAATTTATGACCAAAACCTGAGAGAAATGGGCTGGAGAAAGTTTAC 437  
 QY 241 CTGAGGCAAGTGAAGATACCCCACTGGAGAGAGAGAGAAATCAGCTGTGATCTGGAG 300  
 DB 438 CTGAGGCAAGTGAAGATACCCCACTGGAGAGAGAGAGAAATCAGCTGTGATCTGGAG 497  
 QY 301 CCAAGAAATTCATAGATAGCCATCTCTGGCTTTGGCAGCAGCAGATTTGGGACTCTCCAGAA 360  
 DB 498 CCAAGAAATTCATAGATAGCCATCTCTGGCTTTGGCAGCAGCAGATTTGGGACTCTCCAGAA 557  
 QY 361 GGCATTCAGCAGAAAGTTCTGGTGTGACCTCTTTCAGTGAATGCAAGAGAGGGGCTCA 420  
 DB 558 GGCATTCAGCAGAAAGTTCTGGTGTGACCTCTTTCAGTGAATGCAAGAGAGGGGCTCA 617  
 QY 421 GAAAGCAAGAGGAAATGTTGTTTATTAACCACTTACATCACTCAAGAGAGGCTG 480  
 DB 618 GAAAGCAAGAGGAAATGTTGTTTATTAACCACTTACATCACTCAAGAGAGGCTG 677  
 QY 481 CAATACCGAAGCAGAGGGGCGGTGAGAGCTGCCAAGTGGGGCTTTGGCATCTCTCAT 540  
 DB 678 CAATACCGAAGCAGAGGGGCGGTGAGAGCTGCCAAGTGGGGCTTTGGCATCTCTCAT 737  
 QY 541 CGATCCGCGGCTCTCTCTCATCTACAGCTCTCAACAGATTTCCAGAAATACAGAGAT 600  
 DB 738 CGATCCGCGGCTCTCTCTCATCTACAGCTCTCAACAGATTTCCAGAAATACAGAGAT 797  
 QY 601 GGCCTGCCCAAAATTTCCAAACAGCTGTATTAAGGTGAGAGATGCAAAATGATGTCAAGA 660  
 DB 798 GGCCTGCCCAAAATTTCCAAACAGCTGTATTAAGGTGAGAGATGCAAAATGATGTCAAGA 857  
 QY 661 ATGGCTTCTCAT-GGATCAAAATTTGTCATTCAGCTAAAGATGGGGGCAAAAGACTTACC 718  
 DB 858 ATGGCTTCTCATGGGATCAAAATTTGTCATTCAGCTAAAGATGGGGGCAAAAGACTTACC 916  
  
 RESULT 14  
 CDS19131 813 bp mRNA linear EST 06-JUN-2003  
 LOCUS AGENCOURT 14372790 NIH MGC 181 Homo sapiens cDNA clone  
 DEFINITION IMAGE:30400135 5', mRNA sequence.  
 ACCESSION CDS19131  
 VERSION CDS19131.1 GI:31450849  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
 1 (bases 1 to 813)  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgsdbs-rc@mail.nih.gov  
 Tissue Procurement: Dr. Michael Brownstein  
 cDNA Library Preparation: Invitrogen Corp  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov

FEATURES  
 SOURCE  
 1. 813  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30400135"  
 /issue\_type="White Matter"  
 /dev\_stage="Unknown"  
 /lab\_host="DH10B-Ton A (T1 and T5 phage resistance)"  
 /clone\_id="NIH MGC 181"  
 /note="Vector: pCMV-Sport6.1; Site 1: NotI; Site 2: EcoRV  
 (destroyed); Library is oligo-dT primed and directionally  
 cloned (EcoRV site is destroyed upon cloning). Average  
 insert size 1.42 kb. Library was constructed by  
 (Invitrogen). Note: this is a NIH MGC Library."  
 BASE COUNT 240 a 175 c 218 g 177 t 3 others  
 ORIGIN  
 Query Match 50.8%; Score 682.4; DB 14; Length 813;  
 Best Local Similarity 98.2%; Pred. No. 2e-193;  
 Matches 700; Conservative 0; Mismatches 12; Indels 1; Gaps 1;  
 QY 11 TCTCTAAGAGACTTTTGAAGAAATTAAGAGAAATAGCAGCTGAGATGTTGCTA 70  
 DB 61 TCTAAGAGAGACTTTTGAAGAAATTAAGAGAAATAGCAGCTGAGATGTTGCTA 120  
 QY 71 AAGCAATCATCAACTAGCTGTTATGTTAAAGCCAGAAACAGATCTATGAGGATTTG 130  
 DB 121 AAGCAATCATCAACTAGCTGTTATGTTAAAGCCAGAAACAGATCTATGAGGATTTG 180  
 QY 131 CACTTCTGTTGATCTGTTGAGCCAGACTGAGTGGCTCCAGAACTTGAAAAAGCCA 190  
 DB 181 CACTTCTGTTGATCTGTTGAGCCAGACTGAGTGGCTCCAGAACTTGAAAAAGCCA 240  
 QY 191 TCCAAATTTATGTAACCAAAACCTGAGAGAGTGGGCTGAGAAAGTTCACTGGAGCCAG 250  
 DB 241 TCCAAATTTATGTAACCAAAACCTGAGAGAGTGGGCTGAGAAAGTTCACTGGAGCCAG 300  
 QY 251 TGAAGATACCCCACTGGAGAGAGGAGAGAAATCAGCTGTGATGCTGGAGCCAGAAATTC 310  
 DB 301 TGAAGATACCCCACTGGAGAGAGGAGAGAAATCAGCTGTGATGCTGGAGCCAGAAATTC 360  
 QY 311 ATAAAGATAGCCATCTGAGTCTTGGCGACAGCATTTGGGACTCTCCAGAAAGCAATTAC 370  
 DB 361 ATAAAGATAGCCATCTGAGTCTTGGCGACAGCATTTGGGACTCTCCAGAAAGCAATTAC 420  
 QY 371 CAGAGATTCGAGGAGTACCTCTTTGATGAATGAGAGAGAGGCTCCAGAAAGCAAGAG 430  
 DB 421 CAGAGATTCGAGGAGTACCTCTTTGATGAATGAGAGAGAGGCTCCAGAAAGCAAGAG 480  
 QY 431 GGAAGATTTGTTTATTAACCAACTTATCACTCACTCAAGAGCGGTGCAATACCGAA 490  
 DB 481 GGAAGATTTGTTTATTAACCAACTTATCACTCACTCAAGAGCGGTGCAATACCGAA 540  
 QY 491 CGCAGGGGGCGGTGAGAGCTGCCAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGG 550  
 DB 541 CGCAGGGGGCGGTGAGAGCTGCCAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGG 600  
 QY 551 CTTCTCTCTCATCTAAGTCCACACAGGTATTCAGAAATACAGAGATGGCGGTGCCA 610  
 DB 601 CTTCTCTCTCATCTAAGTCCACACAGGTATTCAGAAATACAGAGATGGCGGTGCCA 660  
 QY 611 AAATTCAGAGCTGTATTAACGTGAGAGATGAGAGATGATGTCAGAAATGGCTTCTC 670  
 DB 661 AGATTCAGAGCTGTATTAACGTGAGAGATGAGAGATGATGTCAGAAATGGCTTCTC 720  
 QY 671 ATGGGATCAAAATTTGTCATTCAGTAAAGAT-GGGGCAAAAGACTTACCCAGA 722  
 DB 721 ATGGGATCAAAATTTGTCATTCAGTAAAGATGGGGGCAAAAGACTTACCCAGA 773



RESULT 15  
 LOCUS CD366462/c 708 bp mRNA linear EST 29-MAY-2003  
 DEFINITION UI-H-FT1-bjt-p-03-0-UI.s1 NCI CGAP\_FTI Homo sapiens cDNA clone  
 CD366462  
 ACCESSION CD366462.1 GI:31150552  
 VERSION EST.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 708)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Dr. Gary W. Hunninghake, U of I  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
 http://genome.uiowa.edu/distribution/cgap.html  
 Seq primer: M13 FORWARD  
 Polya=Yes.

FEATURES  
 source  
 1..708  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-H-FT1-bjt-p-03-0-UI"  
 /tissue\_type="Aveolar Macrophage"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP FTI"  
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
 NCI CGAP\_FTI is a normalized cDNA library constructed from  
 a pool of 81 RNA samples from Alveolar Macrophages  
 challenged with different treatments. The library was  
 normalized according to Bonaldo, Lennon and Soares, Genome  
 Research, 6:791-806, 1996. First strand cDNA synthesis was  
 primed with an oligo-dT primer containing a Not I site.  
 Double stranded cDNA was ligated to an EcoR I adaptor,  
 digested with Not I, and cloned directionally into  
 pT73-Pac vector. The oligonucleotide used to prime the  
 synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (dT)18 tail. The sequence tag for this library is  
 GGCCTATGCCG. The tissue was provided by Dr. Gary W.  
 Hunninghake of the University of Iowa.  
 TAG LIB=UI-H-FT1  
 TAG TISSUE=Human Lung Aveolar Macrophage  
 TAG\_SEQ=GGCCTATGCCG"

BASE COUNT 158 a 192 c 149 g 209 t  
 ORIGIN

Query Match 50.7%; Score 681.8; DB 14; Length 708;  
 Best Local Similarity 99.7%; Pred. No. 2,8e-193;  
 Matches 663; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 206 AAAACCTGAGAGATGGGCTGAGAAAGTTACCTGAGGCACTGAGAAATACCCACT 265  
 DB 708 AAAACCTGAGAGATGGGCTGAGAAAGTTACCTGAGGCACTGAGAAATACCCACT 649  
 QY 266 GGGAGAGGGAGAAATCACTGTGATGCTGAGACCAAGAAATTCATAAGATAGCCATCC 325  
 DB 648 GGGAGAGGGAGAAATCACTGTGATGCTGAGACCAAGAAATTCATAAGATAGCCATCC 589  
 QY 326 TGGGTCTTGAGCAGCATTTGGGACTCTCCAGAAAGCATTTACAGAGAAATTTGGTGG 385

DB 588 TGGGTCTTGAGCAGCATTTGGGACTCTCCAGAAAGCATTTACAGAGAAATTTGGTGG 529  
 QY 386 TGACCTCTTTTGATGAATCTGAGAGAAAGGCTCTAGAAAGCAAGAGGAATTTGTTT 445  
 DB 528 TGACCTCTTTTGATGAATCTGAGAGAAAGGCTCTAGAAAGCAAGAGGAATTTGTTT 469  
 QY 446 ATTAACAACCTTATCATCAATCACTCAAGAGAGGTCATACCGAAGCAAGAGGAGGCTGG 505  
 DB 468 ATTAACAACCTTATCATCAATCACTCAAGAGAGGTCATACCGAAGCAAGAGGAGGCTGG 409  
 QY 506 AAGCTGCCAAGGTGGGGCTTTGGCATCTCTCATTTGATCCGTGGCTCTTCTCAATCT 565  
 DB 408 AAGCTGCCAAGGTGGGGCTTTGGCATCTCTCATTTGATCCGTGGCTCTTCTCAATCT 349  
 QY 566 ACACTCTCAACAGGATTTTCAGAAATTCAGAGATGCGTGGCCCAAAATTCACACGCT 625  
 DB 348 ACACTCTCAACAGGATTTTCAGAAATTCAGAGATGCGTGGCCCAAAATTCACACGCT 289  
 QY 626 GTATTACGGTGGAGATGCAAGAAATGATGTCAGAAATGAGTTCATGGGATCAAAATTTG 685  
 DB 288 GTATTACGGTGGAGATGCAAGAAATGATGTCAGAAATGAGTTCATGGGATCAAAATTTG 229  
 QY 686 TCATTCAAGCTTAAGATGGGGCAAAAGCTTACCAGATGATGATTCCTTCAACACTGTAG 745  
 DB 228 TCATTCAAGCTTAAGATGGGGCAAAAGCTTACCAGATGATGATTCCTTCAACACTGTAG 169  
 QY 746 CAGAGATCATCTGGAGCAAAATATCCAGAACAGTTGTATCTGTAGTGGACATCTGGACA 805  
 DB 168 CAGAGATCATCTGGAGCAAAATATCCAGAACAGTTGTATCTGTAGTGGACATCTGGACA 109  
 QY 806 GCTGGAGATGTTGGCAGGGTGGCATGATGATGAGCGGTGGAGCCTTATATCATGGGAAG 865  
 DB 108 GCTGGAGATGTTGGCAGGGTGGCATGATGATGAGCGGTGGAGCCTTATATCATGGGAAG 49  
 QY 866 CACTCTCACTTATTAAGATCTTGG 890  
 DB 48 CACTCTCACTTATTAAGATCTTGG 24

Search completed: December 22, 2003, 23:11:30  
 Job time : 2632.89 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:59:40 ; Search time 88.7694 Seconds  
(without alignments)  
6682.695 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_171\_1514  
Perfect score: 1344  
Sequence: 1 AAGAAATGCGCTCTTAAGAG.....AAGAAATGCTGCTAGGTCC 1344

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1340.8	99.8	1863	US-09-482-273-28	Sequence 28, Appl
2	772.4	57.5	1134	US-09-482-273-95	Sequence 95, Appl
3	48.2	3.6	7218	US-08-232-463-14	Sequence 14, Appl
4	37	2.8	2049	US-08-268-797-1	Sequence 1, Appl
5	37	2.8	2049	PCT-US95-08414-1	Sequence 1, Appl
6	37	2.8	2691	US-07-878-960-1	Sequence 1, Appl
7	35.2	2.6	2373	US-08-980-080-1	Sequence 1, Appl
8	35.2	2.6	3003	US-08-434-730-15	Sequence 15, Appl
9	34.6	2.6	4086	US-09-220-132-56	Sequence 56, Appl
10	34.6	2.6	4334	US-09-620-312D-88	Sequence 88, Appl
11	34.4	2.6	738	US-08-738-462-1	Sequence 1, Appl
12	34.4	2.6	738	PCT-US94-07587-1	Sequence 1, Appl
13	34	2.5	2106	US-09-252-991A-7477	Sequence 7477, Ap
14	34	2.5	2182	US-09-252-991A-7334	Sequence 7334, Ap
15	34	2.5	2187	US-09-252-991A-7167	Sequence 7167, Ap
16	33.4	2.5	2201	US-09-330-970-2	Sequence 2, Appl
17	33.4	2.5	2201	US-09-330-970-4	Sequence 4, Appl
18	33.2	2.5	2443	US-08-745-934-2	Sequence 2, Appl
19	33.2	2.5	9763	US-08-973-273-1	Sequence 1, Appl
20	33	2.5	1165	US-09-023-942A-28	Sequence 28, Appl
21	33	2.5	1443	US-09-328-352-2934	Sequence 2934, Ap
22	33	2.5	1491	US-09-328-352-1416	Sequence 1416, Ap
23	32.2	2.4	2272	US-09-313-394A-3342	Sequence 3342, Ap
24	32.2	2.4	2220	US-08-997-251-1	Sequence 1, Appl
25	31.8	2.4	4631	US-09-620-312D-164	Sequence 164, Appl
26	31.8	2.4	615	US-09-107-532A-1080	Sequence 1080, Ap
27	31.6	2.4	949	US-09-221-017B-594	Sequence 594, App

28	31.4	2.3	2242	3	US-09-400-742-1	Sequence 1, Appl
29	31.4	2.3	2242	3	US-08-618-651A-1	Sequence 1, Appl
30	31.4	2.3	2242	3	US-09-215-252-1	Sequence 1, Appl
31	31.2	2.3	1712	3	US-08-809-326A-8	Sequence 8, Appl
32	31.2	2.3	1712	4	US-09-689-914A-8	Sequence 8, Appl
33	31.2	2.3	1712	4	US-09-689-913A-8	Sequence 8, Appl
34	31.2	2.3	1712	4	US-09-689-916A-8	Sequence 8, Appl
35	31.2	2.3	1927	4	US-09-689-914A-13	Sequence 13, Appl
36	31.2	2.3	1927	4	US-09-689-913A-13	Sequence 13, Appl
37	31.2	2.3	1927	4	US-09-689-916A-13	Sequence 13, Appl
38	31.2	2.3	1954	3	US-08-809-326A-13	Sequence 13, Appl
39	31	2.3	280	3	US-08-783-853A-34	Sequence 34, Appl
40	31	2.3	280	3	US-09-344-050-34	Sequence 34, Appl
41	31	2.3	321	3	US-08-783-853A-43	Sequence 43, Appl
42	31	2.3	321	4	US-09-344-050-43	Sequence 43, Appl
43	31	2.3	1198	3	US-09-319-989-1	Sequence 1, Appl
44	31	2.3	162450	4	US-09-345-882-1	Sequence 1, Appl
45	30.8	2.3	1044	4	US-09-220-132-146	Sequence 146, App

#### ALIGNMENTS

RESULT 1  
US-09-482-273-28  
Sequence 28, Application US/09482273  
Patent No. 6514631

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 71 Human Secreted Proteins

FILE REFERENCE: P2030P1

CURRENT APPLICATION NUMBER: US/09/482,273

EARLIER FILING DATE: 2000-01-13

EARLIER APPLICATION NUMBER: PCT/US99/15849

EARLIER FILING DATE: 1998-07-15

EARLIER APPLICATION NUMBER: 60/092,922

EARLIER FILING DATE: 1998-07-15

EARLIER APPLICATION NUMBER: 60/092,956

EARLIER FILING DATE: 1998-07-15

NUMBER OF SEQ ID NOS: 267

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 28

LENGTH: 1863

TYPE: DNA

ORGANISM: Homo sapiens

US-09-482-273-28

Query Match

Best Local Similarity 99.8%; Score 1340.8; DB 4; Length 1863;

Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	AAGATGCGCTCTTAAGAGACTTTGAAGAAATTAAGAAATAGCAGCTGGA	60
DB	171	AAGATGCGCTCTTAAGAGACTTTGAAGAAATTAAGAAATAGCAGCTGGA	230
QY	61	GATGTCCTTAAGCATCAACTTGAAGCTTTTATGTTAAGCCGAAACAGATCTAT	120
DB	231	GATGTCCTTAAGCATCAACTTGAAGCTTTTATGTTAAGCCGAAACAGATCTAT	290
QY	121	GAGGATGCGCTCTTGGTTGATCTGTGAGCCAGACTGAGTCTCCAGAACTTA	180
DB	291	GAGGATGCGCTCTTGGTTGATCTGTGAGCCAGACTGAGTCTCCAGAACTTA	350
QY	181	GAAGAGCCATCCAAATTTATGTACCAAACTTCAGCAAGATGGCTGAGAAATTCAC	240
DB	351	GAAGAGCCATCCAAATTTATGTACCAAACTTCAGCAAGATGGCTGAGAAATTCAC	410
QY	241	CTGAGGCGATGAGATATCCCTGAGGAGGAGGAGAGATCAGCTGTGATGCTGAG	300
DB	411	CTGAGGCGATGAGATATCCCTGAGGAGGAGGAGAGATCAGCTGTGATGCTGAG	470

301 CCAAGATTCATTAAGATAGCATCTGGGCTTTGGCAGCAGCATGGGACTCTCCAGAA 360  
471 CCAAGATTCATTAAGATAGCATCTGGGCTTTGGCAGCAGCATGGGACTCTCCAGAA 530  
361 GGCATTACAGCAGAGATTCTGGTGTGACCTCTTTGATGAACTGCAAGAAAGGCTTCA 420  
531 GGCATTACAGCAGAGATTCTGGTGTGACCTCTTTGATGAACTGCAAGAAAGGCTTCA 590  
421 GAAGCAAGAGGGAATGTTGTTATTAACCACTTCACTCACTCACTCACTCACTCACT 480  
591 GAAGCAAGAGGGAATGTTGTTATTAACCACTTCACTCACTCACTCACTCACTCACT 650  
481 CAATCCCAAGCAGAGGAGGCGGTGAGAGCTGCAAGTGGGCGCTTGGCATCTCTCAT 540  
651 CAATCCCAAGCAGAGGAGGCGGTGAGAGCTGCAAGTGGGCGCTTGGCATCTCTCAT 710  
541 CGATCCGTTGCT 600  
711 CGATCCGTTGCT 770  
601 GCGCGGCCCAAAATTCACAGCCCTGATTAAGGTTGGAAGATGCAAGATGATGTCAGA 660  
771 GCGCGGCCCAAAATTCACAGCCCTGATTAAGGTTGGAAGATGCAAGATGATGTCAGA 830  
661 ATGCTCTCTCATGGGATCAAAATTTGCTCATTCAGTAAAGATGGGGCAAAAGCTTAC 720  
831 ATGCTCTCTCATGGGATCAAAATTTGCTCATTCAGTAAAGATGGGGCAAAAGCTTAC 890  
721 GATCTGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATTCAGAACAGTT 780  
891 GATCTGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATTCAGAACAGTT 950  
781 GATCTGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATTCAGAACAGTT 840  
951 GATCTGATTCCTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATTCAGAACAGTT 1010  
841 GGTGAGGCTTTATATCATGGAAGCACTCTCATTTAAAGATCTTGGGCTGCTTCA 900  
1011 GGTGAGGCTTTATATCATGGAAGCACTCTCATTTAAAGATCTTGGGCTGCTTCA 1070  
901 AAGAGACTCTGCGGCTGCTGCTGCACTGCAAGAAACAAGGTGAGTTGCTTCC 960  
1071 AAGAGACTCTGCGGCTGCTGCTGCACTGCAAGAAACAAGGTGAGTTGCTTCC 1130  
961 CAGTATTCATGATTAACAAGGTAATTTTCCACTTACAGTCTGATGAGCTTAC 1020  
1131 CAGTATTCATGATTAACAAGGTAATTTTCCACTTACAGTCTGATGAGCTTAC 1190  
1021 GCAGGAACCTTCTTACCACTGAGGCTGCAATTCATCTGGCAGTAAAGGCCAGGCCATC 1080  
1191 GCAGGAACCTTCTTACCACTGAGGCTGCAATTCATCTGGCAGTAAAGGCCAGGCCATC 1250  
1081 ATGAGAGAGTTTATGAGCTGCTGAGCCCTCAATATCACTCAGGTCTGAGCCATGGA 1140  
1251 ATGAGAGAGTTTATGAGCTGCTGAGCCCTCAATATCACTCAGGTCTGAGCCATGGA 1310  
1141 GAAGGAGAGATCACTTTTGGATCCAAGTGGAGTCCCTGAGGCCAGTCTTATAT 1200  
1311 GAAGGAGAGATCACTTTTGGATCCAAGTGGAGTCCCTGAGGCCAGTCTTATAT 1370  
1201 GACTTATACAGTATTTCTTCTTCATCACTCCACGAGACACATACCTGATGAT 1260  
1371 GACTTATACAGTATTTCTTCTTCATCACTCCACGAGACACATACCTGATGAT 1430  
1261 CCAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC 1320  
1431 CCAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC 1490  
1321 ATGAGAAATGCTGCTGAGTCC 1344  
1491 ATGAGAAATGCTGCTGAGTCC 1514

RESULT 2  
US-09-482-273-95  
; Sequence 95 Application US/09482273  
; Patent No. 6534631  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 71 Human Secreted Proteins  
; FILE REFERENCE: P2030P1  
; CURRENT APPLICATION NUMBER: US/09/482, 273  
; EARLIER FILING DATE: 2000-01-13  
; EARLIER APPLICATION NUMBER: PCT/US99/15849  
; EARLIER FILING DATE: 1999-07-14  
; EARLIER APPLICATION NUMBER: 60/092, 921  
; EARLIER FILING DATE: 1998-07-15  
; EARLIER APPLICATION NUMBER: 60/092, 922  
; EARLIER FILING DATE: 1998-07-15  
; EARLIER APPLICATION NUMBER: 60/092, 956  
; EARLIER FILING DATE: 1998-07-15  
; NUMBER OF SEQ ID NOS: 267  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 95  
; LENGTH: 1134  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-482-273-95  
Query Match 57.5%; Score 772.4; DB 4; Length 1134;  
Best Local Similarity 99.7%; Pred. No. 6e-234;  
Matches 784; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
OY 559 TCCATCTACAGTCTCTCAACAGGATTCAGAAATACAGAGATGGCTGCCAAATTC 618  
DB 1 TCCATCTACAGTCTCTCAACAGGATTCAGAAATACAGAGATGGCTGCCAAATTC 60  
OY 619 ACACCTGTATTAACGTGGAAGATGCAAAATGATGCAAAATGCTTCTCATGGATC 678  
DB 61 ACACCTGTATTAACGTGGAAGATGCAAAATGATGCAAAATGCTTCTCATGGATC 120  
OY 679 AAAATTTCTACGCTTAAGATGGGGCAAAAGCTTACCCAGATCTTCTTCAAC 738  
DB 121 AAAATTTCTACGCTTAAGATGGGGCAAAAGCTTACCCAGATCTTCTTCAAC 180  
OY 739 ACTGTACAGAGATCACTGGGAGCAATATCCAGAAAGATGTTGATCTGTCAGTGA 798  
DB 181 ACTGTACAGAGATCACTGGGAGCAATATCCAGAAAGATGTTGATCTGTCAGTGA 240  
OY 799 CTGACAGCTGGATGTTGGGCAAGGTCATGATGATGATGATGATGATGATGATGAT 858  
DB 241 CTGACAGCTGGATGTTGGGCAAGGTCATGATGATGATGATGATGATGATGATGAT 300  
OY 859 TGGGACCATCTCTTATTAAGATCTTGGGCTGCTGCTCAAAAGAGATCTTGGGCTG 918  
DB 301 TGGGACCATCTCTTATTAAGATCTTGGGCTGCTGCTCAAAAGAGATCTTGGGCTG 360  
OY 919 GTGCTCTGACCTGCAAGAAACAAGGTGAGTGGCTTCTTCCAGTATTAAGTTAC 978  
DB 361 GTGCTCTGACCTGCAAGAAACAAGGTGAGTGGCTTCTTCCAGTATTAAGTTAC 420  
OY 979 AAGTAAATATTTTCAACTACAGTCTGATGATGATGATGATGATGATGATGATGAT 1038  
DB 421 AAGTAAATATTTTCAACTACAGTCTGATGATGATGATGATGATGATGATGATGAT 480  
OY 1039 ACTGGGCTGCAATCACTGGCAGTGAAGAGCCAGGCCCATATGAGAGGTTATAGC 1098  
DB 481 ACTGGGCTGCAATCACTGGCAGTGAAGAGCCAGGCCCATATGAGAGGTTATAGC 539  
OY 1099 CTGCTGAGCCCTCAATATCACTCAGGCTCTGAGCCATGAGAGAGAGAGAGATCAAC 1158  
DB 540 CTGCTGAGCCCTCAATATCACTCAGGCTCTGAGCCATGAGAGAGAGAGATCAAC 599  
OY 1159 TTTTGGATCCAAAGCTGAGTGGCTGAGCCAGTCTTATGATGATGATGATGATGAT 1218  
DB 600 TTTTGGATCCAAAGCTGAGTGGCTGAGCCAGTCTTATGATGATGATGATGATGAT 659

Qy 1219 TTCTTCATCACTCCACGAGACACATGATGATGATCCAAAGCAGATGATGTT 1278  
Db 660 TTCTTCATCACTCCACGAGACACATGATGATGATCCAAAGCAGATGATGTT 719  
Qy 1279 GCTGCTGCTGTTGGGCTGTTGTTTCTTATGTTGTCAGACATGAGAAATGCTGCT 1338  
Db 720 GCTGCTGCTGTTGGGCTGTTGTTTCTTATGTTGTCAGACATGAGAAATGCTGCT 779  
Qy 1339 AGTCC 1344  
Db 780 AGTCC 785

RESULT 3  
US-08-232-463-14/c  
Sequence 14, Application US/08232463  
Patent No. 3670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-9300  
TELEFAX: (703) 683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZSpt-Fls  
US-08-232-463-14

Query Match 3.6%; Score 48.2; DB 1; Length 7218;  
Best Local Similarity 2.8%; Pred. No. 0.00018;  
Matches 11; Conservative 220; Mismatches 158; Indels 0; Gaps 0;  
Qy 149 TTGACCCAGACAGTGTGCTCCAAAGACTGAAAGCAATTAATGATCAAA 208  
Db 1442 TTTGGTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1383  
Qy 209 ACTGCAGCAAGATGGGTGAGAAAGTTCACTGAGACCATGAGAAATACCCACTGG 268  
Db 1300 GCTCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1841

Db 1382 RRR 1323  
Qy 269 AGAGGGAGAGATGAGCTGTGATGTCGAGCCAAAGATTCATAAGATACCATCTCG 328  
Db 1332 RRR 1263  
Qy 329 GTCTTGACGACGATTCCTCTCCAGAGCATTAACAGAACTGCTGCTGTA 388  
Db 1262 RRR 1203  
Qy 389 CCTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 448  
Db 1202 RRR 1143  
Qy 449 ACCAAGCTTACATGATGATGATGATGATGATGATGATGATGATGATGAT 508  
Db 1142 RRR 1083  
Qy 509 CTGCAAGTGGGGCTTGGCATCTCTC 537  
Db 1082 RRR 1054

RESULT 4  
US-08-268-797-1/c  
Sequence 1, Application US/08268797  
Patent No. 5597788  
GENERAL INFORMATION:  
APPLICANT: Purchio, Anthony F.  
APPLICANT: LeBaron, Richard  
TITLE OF INVENTION: Factor to Grow Tissue Ex Vivo  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson and Bear  
STREET: 620 Newport Center Drive, Sixteenth Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/268,797  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: TISSUE.001A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-0176  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2049 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-268-797-1

Query Match 2.8%; Score 37; DB 1; Length 2049;  
Best Local Similarity 49.7%; Pred. No. 0.29;  
Matches 94; Conservative 0; Mismatches 95; Indels 0; Gaps 0;  
Qy 1125 GGTCTGAGCCATGAGAGAGGACAGACATCACTTTGATCCAGCTGAGTCTCG 1184  
Db 1300 GCTCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1841

Query Match	2.8%	Score 37	DB 5	Length 2049
Best Local Similarity	49.7%	Pred. No. 0.29		
Matches	94	Conservative	0	Mismatches 95; Indels 0; Gaps 0;
QY	1125	GGTCTTAGCCATGAGAGGAGACAGACATCAATCTTTGGATGCCAAGCTGAGTGCCTCG	1184	
Db	1900	GCTCGAGAACTTGGTGATGACATGGACCAAGCCATTTGGCCATGATGTGCAAGCTCG	1841	
QY	1185	AGCCAGTCTACTTGATGACTTAATACAGTATTTCTTCCATCACTGCCAGGAGAC	1244	
Db	1840	CACAGGCTCTTGTGTGACACTCACCAATGTTTTTCAGCTGACTTCCAGCTGTGTAC	1781	
QY	1245	CATCACTGTACGTGCCAAGAGAGATGAATGTGCTGCTGTTGGCTGTGTTC	1304	

[illegible]

Db 1827 CTGGAGAGACTTACCGGACCAAGGCGCCGATGCTCGGCTAACGAGATTTCATCAC 1768  
 QY 1305 TTATGTGT 1313  
 Db 1767 CAATGTGT 1759

RESULT 7

US-08-980-080-1  
 : Sequence 1, Application US/08980080  
 : Patent No. 6312941  
 : GENERAL INFORMATION:  
 : APPLICANT: CARTER-SU, CHRISTIN  
 : APPLICANT: RUI, LIANG-YOU  
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING  
 : TITLE OF INVENTION: SIGNALING PATHWAY AGONISTS AND ANTAGONISTS  
 : NUMBER OF SEQUENCES: 6  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: MEDLEN & CARROLL, LLP  
 : STREET: 220 MONTGOMERY STREET, SUITE 2200  
 : CITY: SAN FRANCISCO  
 : STATE: CALIFORNIA  
 : COUNTRY: UNITED STATES OF AMERICA  
 : ZIP: 94104  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patentin Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/980,080  
 : FILING DATE: 26-NOV-1997  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: CARROLL, PETER G.  
 : REGISTRATION NUMBER: 32,837  
 : REFERENCE/DOCKET NUMBER: UM-03036  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (415) 705-8410  
 : TELEFAX: (415) 397-8338  
 : INFORMATION FOR SEQ ID NO: 1:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 2373 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: double  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: DNA (genomic)  
 : US-08-980-080-1

Query Match 2.6%; Score 35.2; DB 4; Length 2373;

Best Local Similarity 44.4%; Pred. No. 1.2; Mismatches 178; Indels 0; Gaps 0;

Matches 142; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 771 AGACAGGTTGATGTCAGTGCATCTGACAGCTGGAGTGTGGGACGGTGCAT 830  
 Db 729 AGCAGAGTGTATACAGAGAGAGCTGCTGATTTATGGGGGTGAAGGCTGCC 788  
 QY 831 GGATGATGGCGGTGAGCTTTATATCATGGGAAGCACTTCACTTATTAAGCTTGG 890  
 Db 789 TGACCCAGCAGAGATGCTGCTGGAGAGGGGCAAGCTGGCTGAACCTCGAGAGGAGG 848  
 QY 891 GCTGCTCAAGAGAGAGCTGCGGCTGCTGCTGCACTGCAAGAAAGCAAGGTGAGT 950  
 Db 849 GCAGCTCAGTGGCAGAAATGTGATTAAGCTGCTCGGAGTGAAGAGAGAGAGG 908  
 QY 951 TGTGCTTCCAGTATTAATACATACAGAGTAATATTTCCAACTACAGTCTGAT 1010  
 Db 909 AAGTGTGATGAGTCTTTGATACACCAAGGCAATCCGGCCGCTTTAGCATTCCTG 968  
 QY 1011 GGAAGTCAAGCAGAACTTTTACCACTGGGCTGCAATTCAGTGCAGTGAAGGC 1070  
 Db 969 TTCTACTATTACTGATGTCGACAGCCAGCCCTGAGATGCTGACAGGAGAACAC 1028

QY 1071 CAGGCCATCATGAGAGAG 1090  
 Db 1029 GTTGTGTTAAGGTAGAG 1048

RESULT 8

US-08-434-730-15  
 : Sequence 15, Application US/08434730  
 : Patent No. 5637463  
 : GENERAL INFORMATION:  
 : APPLICANT: Dalton, Stephen  
 : APPLICANT: Kochan, Jarema P  
 : APPLICANT: Osborne, Mark A  
 : TITLE OF INVENTION: METHOD TO DETECT PROTEIN-PROTEIN  
 : TITLE OF INVENTION: INTERACTIONS  
 : NUMBER OF SEQUENCES: 16  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Hoffmann-La Roche Inc  
 : STREET: 340 Kingsland Street  
 : CITY: Nutley  
 : STATE: NJ  
 : COUNTRY: USA  
 : ZIP: 07110  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patentin Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/434,730  
 : FILING DATE: 04-MAY-1995  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Semionow, Raina  
 : REGISTRATION NUMBER: 39022  
 : REFERENCE/DOCKET NUMBER: 9069  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (201)235-4391  
 : TELEFAX: (201)235-2363  
 : INFORMATION FOR SEQ ID NO: 15:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 3003 base pairs  
 : TYPE: nucleic acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: cDNA  
 : HYPOTHETICAL: YES  
 : ANTI-SENSE: NO  
 : US-08-434-730-15

Query Match 2.6%; Score 35.2; DB 1; Length 3003;

Best Local Similarity 44.4%; Pred. No. 1.3; Mismatches 178; Indels 0; Gaps 0;

Matches 142; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 771 AGACAGGTTGATGTCAGTGCATCTGACAGCTGGAGTGTGGGACGGTGCAT 830  
 Db 1072 AGCAGAGTGTATACAGAGAGAGCTGCTGATTTATGGGGGTGAAGGCTGCC 1131  
 QY 831 GGATGATGGCGGTGAGCTTTATATCATGGGAAGCACTTCACTTATTAAGCTTGG 890  
 Db 1132 TGACCCAGCAGAGATGCTGCTGGAGAGGGGCAAGCTGGCTGAACCTCGAGAGG 1191  
 QY 891 GCTGCTCAAGAGAGAGCTGCGGCTGCTGCTGCACTGCAAGAAAGCAAGGTGAGT 950  
 Db 1192 GCAGCTCAGTGGCAGAAATGTGATTAAGCTGCTCGGAGTGAAGAGAGAGAGG 1251  
 QY 951 TGTGCTTCCAGTATTAATACATACAGAGTAATATTTCCAACTACAGTCTGAT 1010  
 Db 1252 AAGTGTGATGAGTCTTTGATACACCAAGGCAATCCGGCCGCTTTAGCATTCCTG 1311  
 QY 1011 GGAAGTCAAGCAGAACTTTTACCACTGGGCTGCAATTCAGTGCAGTGAAGGC 1070  
 Db 1312 TTCTACTATTACTGATGTCGACAGCCAGCCCTGGAATGCTGACAGGAGAACAC 1371

OY 1071 CAGGGCCATCATGAGAGG 1090  
DB 1372 GTTTGTGTAGTAGAG 1391

## RESULT 9

US-09-220-132-56/c  
Sequence 56, Application US/09220132  
Patent No. 6506607

## GENERAL INFORMATION:

APPLICANT: Shyjan, Andrew W.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT

FILE REFERENCE: 07334-074001

CURRENT APPLICATION NUMBER: US/09/220,132

PRIOR FILING DATE: 1998-12-23

PRIOR APPLICATION NUMBER: US 60/079,303

PRIOR FILING DATE: 1998-03-25

NUMBER OF SEQ ID NOS: 191

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 56

LENGTH: 4086

TYPE: DNA

ORGANISM: Homo sapiens

US-09-220-132-56

Query Match 2.6%; Score 34.6; DB 4; Length 4086;

Best Local Similarity 56.6%; Pred. No. 2.5;

Matches 64; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

OY 1179 GCCTGAGCCGCTACTTATGATGATATACAGATTTCTTCCATCACTCCACGG 1238  
DB 1678 GCATGTAGCCCGCTGCTGCTGAACTCACTGACATGATGATCTCCGACCTTCCACT 1619  
DB 1618 CCAAGCTCTTGTCTTCCAGCAGCAATCTCTATGATGATTTTGTAT 1566

## RESULT 10

US-09-620-312D-88/c

Sequence 88, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyun

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yunqing

APPLICANT: Wang, Dunhui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Drmanac, Radojic T.

TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and

FILE REFERENCE: 784CIP2B

CURRENT APPLICATION NUMBER: US/09/620,312D

PRIOR FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: pf\_FL\_genes Version 1.0  
SEQ ID NO 88  
LENGTH: 4334

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (64)..(3810)

US-09-620-312D-88

Query Match 2.6%; Score 34.6; DB 4; Length 4334;

Best Local Similarity 56.6%; Pred. No. 2.6;

Matches 64; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

OY 1179 GCCTGAGCCGCTACTTATGATGATATACAGATTTCTTCCATCACTCCACGG 1238  
DB 1915 GCATGTAGCCCGCTGCTGCTGAACTCACTGACATGATGATCTCCGACCTTCCACT 1856  
OY 1239 AGACACATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1291  
DB 1855 CCAAGCTCTTGTCTTCCAGCAGCAATCTCTATGATGATTTTGTAT 1803

## RESULT 11

US-08-738-462-1/c

Sequence 1, Application US/08738462

Patent No. 5965401

GENERAL INFORMATION:

APPLICANT: Chang, Chiwen

APPLICANT: Lanier, Lewis L.

APPLICANT: Phillips Jr., Joseph H.

TITLE OF INVENTION: Purified Mammalian NK Antigen and

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESS: DNAX Research Institute

STREET: 901 California Avenue

CITY: Palo Alto

STATE: California

COUNTRY: USA

ZIP: 94304-1104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/738,462

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/093,435

FILING DATE: 16-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: Ching, Edwin P.

REGISTRATION NUMBER: 34,090

REFERENCE/DOCKET NUMBER: DX0397

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-496-1200

TELEFAX: 415-852-9196

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 738 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 61..738

US-08-738-462-1

Query Match 2.6%; Score 34.4; DB 2; Length 738;  
Best Local Similarity 48.4%; Pred. No. 1;  
Matches 124; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

QY 624 CTGATTAACGTTGGAAGATGAGAAATGATGTCAAGATGGCTTCTCATGGGATCAAAAT 683  
DB 261 CTGATTAAGAGATGTCACGTGAACATCCAGTAACAACAAAGAGAAATAT 202  
QY 684 TGTTCATTCAGTAAAGATGGGGCAAGACCTACCCAGAT-ACGTATCTTCAACATG 742  
DB 201 CCCAGCACAGCTTAAGTTTCAGGGCAAAATGATGCAAGGTAAACCTGACAGACATCCG 142  
QY 743 TAGCAGAGATCACTGGAGCAAAATATCCAGAAAGTTGTACTGTGATGAGATCTGG 802  
DB 141 AGGAAGAGATGAAGTGAAGAACTTTCTGGGCTGAGTCTGTGGTAAGTTAACTGAC 82  
QY 803 ACAGCTGGAGATGTTGGCAGGGTGCCATGATGATGCGGTGGAGCCTTTATATCATGG 862  
DB 81 ATATATTGCTGTGTGTCATGCGACAGAGAAAGGTGCATTAACTGTGTGTAAGA 22

QY 863 AAGCCTCTCACTTAT 878  
DB 21 ACAAACTCTCAATTCT 6

RESULT 12  
PCT-US94-07587-1/c  
Sequence 1, Application PC/TUS9407587  
GENERAL INFORMATION:  
APPLICANT: Schering Corp.  
TITLE OF INVENTION: PURIFIED MAMMALIAN NK ANTIGENS AND  
TITLE OF INVENTION: RELATED REAGENTS  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Schering Corp.  
STREET: One Giralda Farms  
CITY: Madison  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07940  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh 7.1  
SOFTWARE: Microsoft Word 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/07587  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lunn, Paul G.  
REGISTRATION NUMBER: 32,743  
REFERENCE/DOCKET NUMBER: DX0397K  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-822-7255  
TELEFAX: 201-822-7039  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 738 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURES:  
NAME/KEY: CDS  
LOCATION: 61..738  
PCT-US94-07587-1

Query Match 2.6%; Score 34.4; DB 5; Length 738;  
Best Local Similarity 48.4%; Pred. No. 1;  
Matches 124; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

QY 624 CTGATTAACGTTGGAAGATGAGAAATGATGTCAAGATGGCTTCTCATGGGATCAAAAT 683

DB 261 CTGATTAAGAGATGTCACGTGAACATCCAGTAACAACAAAGAGAAATAT 202  
QY 684 TGTTCATTCAGTAAAGATGGGGCAAGACCTACCCAGAT-ACGTATCTTCAACATG 742  
DB 201 CCCAGCACAGCTTAAGTTTCAGGGCAAAATGATGCAAGGTAAACCTGACAGACATCCG 142  
QY 743 TAGCAGAGATCACTGGAGCAAAATATCCAGAAAGTTGTACTGTGATGAGATCTGG 802  
DB 141 AGGAAGAGATGAAGTGAAGAACTTTCTGGGCTGAGTCTGTGGTAAGTTAACTGAC 82  
QY 803 ACAGCTGGAGATGTTGGCAGGGTGCCATGATGATGCGGTGGAGCCTTTATATCATGG 862  
DB 81 ATATATTGCTGTGTGTCATGCGACAGAGAAAGGTGCATTAACTGTGTGTAAGA 22

QY 863 AAGCCTCTCACTTAT 878  
DB 21 ACAAACTCTCAATTCT 6

RESULT 13  
US-09-252-991A-7477  
Sequence 7477, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
SEQ ID NO 7477  
NUMBER OF SEQ ID NOS: 33142  
LENGTH: 2106  
TYPE: DNA  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7477

Query Match 2.5%; Score 34; DB 4; Length 2106;  
Best Local Similarity 54.9%; Pred. No. 2.6;  
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 970 CAGTTACACAGGTAATATTTCCAACTACAGTCTGTGATGAGTGTGACGACGAACC 1029  
DB 1181 CAGGAAGTGAAGGACTGATGTCTTCTCGCTGTGGTGAACCCGACGACGAAC 1240  
QY 1030 TTCTTACCACTGGCTGCAATTCACTGGCAGTGAAGCCAGGCGCATCATGAGAG 1089  
DB 1241 GCTTCTCAGGGTGAATCAAGTCCGCGCGGAAATCGTTGGCGACCTTGAGAG 1300

QY 1090 GT 1091  
DB 1301 CT 1302

RESULT 14  
US-09-252-991A-7334  
Sequence 7334, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
PRIOR FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7334  
; LENGTH: 2112  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7334

Query Match 2.5%; Score 34; DB 4; Length 2112;  
Best Local Similarity 54.9%; Pred. No. 2.6;  
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 970 CAGTTACACAGGTAAATATTCCACTACAGTCTGTGATGAGTCTGACGACGAACC 1029  
DB 1243 CAGGAAGTGAAGACCTGATCTCTACTTCGCCCTGTGTGTAACCCGACGACGACAC 1302  
QY 1030 TTCTTACCCACTGGGCTGCAATTCACTGCGACGTGAAGGCCAGGGCCATCATGAGAG 1089  
DB 1303 GCCTTCCTCAGGCTGATCAACGTGCGCGCGGAATCGGTTGCGCGACCTGAGAGAG 1362  
QY 1090 GT 1091  
DB 1363 CT 1364

## RESULT 15

US-09-252-991A-7167/c  
; Sequence 7167, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7167  
; LENGTH: 2187  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7167

Query Match 2.5%; Score 34; DB 4; Length 2187;  
Best Local Similarity 54.9%; Pred. No. 2.7;  
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 970 CAGTTACACAGGTAAATATTCCACTACAGTCTGTGATGAGTCTGACGACGAACC 1029  
DB 975 CAGGAAGTGAAGACCTGATCTCTACTTCGCCCTGTGTGTAACCCGACGACGACAC 916  
QY 1030 TTCTTACCCACTGGGCTGCAATTCACTGCGACGTGAAGGCCAGGGCCATCATGAGAG 1089  
DB 915 GCCTTCCTCAGGCTGATCAACGTGCGCGCGGAATCGGTTGCGCGACCTGAGAGAG 856  
QY 1090 GT 1091  
DB 855 CT 854

Search completed: December 22, 2003, 23:17:21  
Job time : 90.7694 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 14:49:05 ; Search time 420.487 Seconds  
(without alignments)  
10650.710 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_171\_1514  
Perfect score: 1344  
Sequence: 1 AAGAAATGCTCTTAAGAG.....AAGAAATGCTCTTAAGTCC 1344

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapept 1.0

Searched: 2211978 seqs, 1666101734 residues  
Total number of hits satisfying chosen parameters: 4423956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- Published Applications NA:\*
- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
  - 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
  - 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
  - 4: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
  - 5: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
  - 6: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
  - 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
  - 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
  - 9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
  - 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
  - 11: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
  - 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
  - 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
  - 14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
  - 15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
  - 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
  - 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
  - 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1344	100.0	1851	9	US-09-745-763-35
2	1340.8	99.8	1863	11	US-09-984-271-28
3	1060.8	78.9	1778	10	US-09-917-800A-505
4	772.4	57.5	1134	11	US-09-984-271-95
5	261.8	19.5	357	10	US-09-833-381-1929
6	215.8	16.1	427	10	US-09-833-381-1930
7	91	6.8	134	10	US-09-783-590-12141
8	60.2	4.5	65	13	US-09-908-975-25931
9	42	3.1	2200	15	US-09-908-975-8960
10	41	3.1	65	13	US-09-908-975-1202
11	39.4	2.9	484	11	US-09-918-995-30889
12	39.4	2.9	1363	15	US-10-232-484-5
13	39.4	2.9	2319	14	US-10-044-090-579
14	39.4	2.9	3322	13	US-09-814-353-20277
15	39.4	2.9			Sequence 20277, A

C 16	38.6	2.9	3997	15	US-10-198-846-12639	Sequence 12639, A
17	38	2.8	457	13	US-10-027-632-310889	Sequence 310889,
18	38	2.8	457	13	US-10-027-632-310889	Sequence 310889,
19	37.8	2.8	999	15	US-10-184-644-434	Sequence 434, App
20	37.8	2.8	999	15	US-10-184-644-434	Sequence 434, App
C 21	37	2.8	502	15	US-10-066-543-2028	Sequence 2028, App
C 22	37	2.8	548	15	US-10-066-543-2222	Sequence 2222, App
C 23	37	2.8	568	15	US-09-998-598-2545	Sequence 2545, App
C 24	37	2.8	2691	9	US-09-932-217-121	Sequence 121, App
C 25	37	2.8	2691	10	US-09-833-263-121	Sequence 121, App
C 26	37	2.8	2691	10	US-09-880-107-2399	Sequence 2399, App
C 27	37	2.8	2691	13	US-10-301-822-200	Sequence 200, App
C 28	37	2.8	2691	13	US-09-873-367C-1009	Sequence 1009, App
C 29	37	2.8	2691	14	US-10-025-380-121	Sequence 121, App
C 30	37	2.8	2691	15	US-10-102-524-1708	Sequence 1708, App
C 31	37	2.8	2782	14	US-10-044-090-647	Sequence 647, App
C 32	37	2.8	2782	15	US-10-208-408-16	Sequence 16, App1
C 33	37	2.8	2782	15	US-10-084-817-28	Sequence 28, App1
C 34	37	2.8	3313	13	US-09-814-353-19947	Sequence 19947, A
C 35	37	2.8	3320	15	US-10-198-846-11549	Sequence 11549, A
C 36	36.4	2.7	5085	15	US-10-198-846-9854	Sequence 9854, App
C 37	35.4	2.6	2574	13	US-10-311-626-13	Sequence 13, App1
C 38	35.2	2.6	1115	15	US-10-184-644-440	Sequence 440, App
C 39	35.2	2.6	1115	15	US-10-184-644-440	Sequence 440, App
C 40	34.6	2.6	416	11	US-09-918-995-16569	Sequence 16569, A
C 41	34.6	2.6	1237	13	US-10-027-632-199318	Sequence 199318,
C 42	34.6	2.6	1237	14	US-10-027-632-199318	Sequence 199318,
C 43	34.6	2.6	4182	15	US-10-060-036-66	Sequence 66, App1
C 44	34.6	2.6	4182	15	US-10-060-036-137	Sequence 137, App
C 45	34.6	2.6	4334	13	US-10-117-722-88	Sequence 88, App1

ALIGNMENTS

RESULT 1  
US-09-745-763-35  
Sequence 35, Application US/09745763  
Patent No. US20020065394A1  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
McCoy, John M.  
Lavallee, Edward R.  
Collins-Racie, Lisa A.  
Evans, Cheryl  
Werberg, David  
Treacy, Maurice  
Spaulding, Vikki  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM  
NUMBER OF SEQUENCES: 219  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/745,763  
FILING DATE: 18-Jun-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Sprunger, Suzanne A.  
REGISTRATION NUMBER: 41,323  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8284  
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1851 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 35:  
US-09-745-763-35

Query Match 100.0%; Score 1344; DB 9; Length 1851;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AAGATGGCATCTCTAAGAGACCTTTTGAAGAAATTAAGAAATGCGCCTGTGGA 60
DB 171 AAGATGGCATCTCTAAGAGACCTTTTGAAGAAATTAAGAAATGCGCCTGTGGA 230
QY 61 GATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAGAACAGATCTAT 120
DB 231 GATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAGAACAGATCTAT 290
QY 121 GAGCGATTGGCACTTCTGTTGATATCTGTGGACCCAGACTGAGTGGCTCCAAGACCTA 180
DB 291 GAGCGATTGGCACTTCTGTTGATATCTGTGGACCCAGACTGAGTGGCTCCAAGACCTA 350
QY 181 GAAAAAGCCATCCAAATTAATGTAACAAAACCTGCGACGAAGATGGGCTGGAAGAAATTGAC 240
DB 351 GAAAAAGCCATCCAAATTAATGTAACAAAACCTGCGACGAAGATGGGCTGGAAGAAATTGAC 410
QY 241 CTGGAGCCAGTGAAGATACCCCACTGGAGAGAGAGAGAAATACAGCTGTGATGTGGAG 300
DB 411 CTGGAGCCAGTGAAGATACCCCACTGGAGAGAGAGAGAAATACAGCTGTGATGTGGAG 470
QY 301 CCAAGAAATTCATAGATAGCCATCTGGGCTTGGACAGACAGATGGGACTCTCCAGAA 360
DB 471 CCAAGAAATTCATAGATAGCCATCTGGGCTTGGACAGACAGATGGGACTCTCCAGAA 530
QY 361 GGCATTACAGCAAGATTTCTGTGTGATCTTTTCGATGAATCTGCAAGAAAGGCGCTCA 420
DB 531 GGCATTACAGCAAGATTTCTGTGTGATCTTTTCGATGAATCTGCAAGAAAGGCGCTCA 590
QY 421 GAAGCAAGAGGGAATGTTGTTTATATACCAACCTTTCATCAATCACTCAAGACGGTG 480
DB 591 GAAGCAAGAGGGAATGTTGTTTATATACCAACCTTTCATCAATCACTCAAGACGGTG 650
QY 481 CAATACCGACGACGAGGGGCGGTGGAAGCTGCCAAGTGGGGGCTTTGGCATCTCTCAT 540
DB 651 CAATACCGACGACGAGGGGCGGTGGAAGCTGCCAAGTGGGGGCTTTGGCATCTCTCAT 710
QY 541 CGATCCGTTGGCTCTTCTCCATCTACAGTCTCTCAACAGGTAATTCAGAAATCCAGAT 600
DB 711 CGATCCGTTGGCTCTTCTCCATCTACAGTCTCTCAACAGGTAATTCAGAAATCCAGAT 770
QY 601 GGGCGGCCCAAAATTCACAGCCTGTATTAAGGTAAGATGCAAGAAATGTCAGAA 660
DB 771 GGGCGGCCCAAAATTCACAGCCTGTATTAAGGTAAGATGCAAGAAATGTCAGAA 830
QY 661 ATGGCTTCTCATGGGATCAAAATTTGTCATTCAGCTAAAGATGGGGGCAAAAGACTTACCA 720
DB 831 ATGGCTTCTCATGGGATCAAAATTTGTCATTCAGCTAAAGATGGGGGCAAAAGACTTACCA 890
QY 721 GATATCTGATTTCTTCAACACTGTAGACAGATCACTGGGAGCAATATTCAGAACAGGTT 780
DB 891 GATATCTGATTTCTTCAACACTGTAGACAGATCACTGGGAGCAATATTCAGAACAGGTT 950
QY 781 GTACGTGTCAGTGCATCTGTCGACGCTGGAGATGTTGGCAGGGGCAATGATGATGGC 840
DB 951 GTACGTGTCAGTGCATCTGTCGACGCTGGAGATGTTGGCAGGGGCAATGATGATGGC 1010
QY 841 GGTGAGGCTTTATATCATGGAAGACACTCTCACTTATTAAGATCTTGGCTGGCTGCA 900
DB 1011 GGTGAGGCTTTATATCATGGAAGACACTCTCACTTATTAAGATCTTGGCTGGCTGCA 1070
```

```

QY 901 AAGAGACCTGCGGGCTGTGCTCTGACTGACAGAAACAAGGTGGAGTTGTCCTTC 960
DB 1071 AAGAGACCTGCGGGCTGTGCTCTGACTGACAGAAACAAGGTGGAGTTGTCCTTC 1130
QY 961 CAGTATTATCAGTTACACAAGATTAATTTTCAACTACAGTCTGTGATGAGTCTGAC 1020
DB 1131 CAGTATTATCAGTTACACAAGATTAATTTTCAACTACAGTCTGTGATGAGTCTGAC 1190
QY 1021 GCAGGAACCTTTTCAACCACTGGGCTGCAATTCATCTGCACTGAAAGGCCAGGCCATC 1080
DB 1191 GCAGGAACCTTTTCAACCACTGGGCTGCAATTCATCTGCACTGAAAGGCCAGGCCATC 1250
QY 1081 ATGAGAGGATTTATAGAGCTGCTGACAGCCCTCAATATCACTGAGTCTGAGCCATGGA 1140
DB 1251 ATGAGAGGATTTATAGAGCTGCTGACAGCCCTCAATATCACTGAGTCTGAGCCATGGA 1310
QY 1141 GAAGGACAGACATCAACTTTTGGATCCAAGCTGAGTGGCTGAGGCCAGTCTTATGAT 1200
DB 1311 GAAGGACAGACATCAACTTTTGGATCCAAGCTGAGTGGCTGAGGCCAGTCTTATGAT 1370
QY 1201 GACTTATACAAATATTTCTTCTTCATCACTCCACGAGACACATGATCTGTATGAT 1260
DB 1371 GACTTATACAAATATTTCTTCTTCATCACTCCACGAGACACATGATCTGTATGAT 1430
QY 1261 CCAAGGACAGATGATGTTGCTGCTGCTGTTGGCTGTGTTTCTTATGTTTGCAGAC 1320
DB 1431 CCAAGGACAGATGATGTTGCTGCTGCTGTTGGCTGTGTTTCTTATGTTTGCAGAC 1490
QY 1321 ATGGAAGAAATGCTGCTAGGTCC 1344
DB 1491 ATGGAAGAAATGCTGCTAGGTCC 1514
```

## RESULT 2

US-09-984-271-28  
Sequence 28, Application US/0984271  
Publication No. US2003004088A1

GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 71 Human Secreted Proteins

FILE REFERENCE: P2030P1  
CURRENT APPLICATION NUMBER: US/09/984,271

PRIOR FILING DATE: 2001-10-29  
CURRENT FILING DATE: 09/482,273

PRIOR APPLICATION NUMBER: 09/482,273  
PRIOR FILING DATE: 2000-01-13

PRIOR APPLICATION NUMBER: PCT/US99/15849  
PRIOR FILING DATE: 1999-07-14

PRIOR APPLICATION NUMBER: 60/092,921  
PRIOR FILING DATE: 1998-07-15

PRIOR APPLICATION NUMBER: 60/092,922  
PRIOR FILING DATE: 1998-07-15

PRIOR APPLICATION NUMBER: 60/092,956  
PRIOR FILING DATE: 1998-07-15

NUMBER OF SEQ ID NOS: 267  
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 28  
LENGTH: 1863

TYPE: DNA  
ORGANISM: Homo sapiens

US-09-984-271-28

Query Match 99.8%; Score 1340.8; DB 11; Length 1863;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1342; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 AAGATGGCATCTCTAAGAGACCTTTTGAAGAAATTAAGAAATGCGCCTGTGGA 60
DB 171 AAGATGGCATCTCTAAGAGACCTTTTGAAGAAATTAAGAAATGCGCCTGTGGA 230
QY 61 GATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAGAACAGATCTAT 120
DB 231 GATGTTGCTAAAGCAATCATCACTAGCTGTTTATGTTAAAGCCAGAACAGATCTAT 290
```



Db 373 GAGAAAGTATCCAAATCATGTATACCAAAACCTGCAACAAGATGGCTGAAAAGCTTCAC 432  
Qy 241 CTGAGCCAGTGAAGATATCCCACTGGAGAGGGGAGAAATACAGCTGTATGTGGAG 300  
Db 433 CTGGAGCGGTGAGAAATCTCTCACTGGGGGAGGGGCGAAATGTGCAGTATGTGTGTG 492  
Qy 301 CCAAGAAATTCATTAAGATAGCCATCTGGGTCTTGGCAGACAGATGGAGACTCTCCAGAA 360  
Db 493 CCTGAATTCACAAATTTGGCTATTTTAGGCTTGGCGGAGCATTTGGAGCTCTCTGAA 552  
Qy 361 GGCATTACAGCAGAAAGTTCTGGTGTGACCTCTTTCATGAACTGACAGAGAGGGCTCA 420  
Db 553 GGTATACAGCAGAAAGTACTGGTGTGGCTCTTTTGTGAACCTTCAAAAGAGGCAATCA 612  
Qy 421 GAAGCAAGAGGAGAGATTTGTTTATTAACCACTTACATCACTACTCAAGAGCGGTG 480  
Db 613 GAGGCAAGAGGAGAAATTTGTTTATTAACCACTTACATCACTACTCAAGAGCGGTG 672  
Qy 481 CAATACCGAAGCGAGGGGGCGGTGAGCTGCAAGTGGGGGCTTTGGCATCTCTCATTT 540  
Db 673 CAGTACCGGAGCGGAGCTGTGAGCTGCAAGTGGGGGCGGTGCAATCTCTCATTC 732  
Qy 541 CGATCCGTGGCTCTCTCTCATCTACAGTCTCTCAACAGGATTTCAAGAAATACAGAGAT 600  
Db 733 CGATCAGTAGCTTCTTTTTCATCTACAGTCTCTCAACAGGATTTCAAGAAATACAGAT 792  
Qy 601 GCGGTGCCCAAAATTCACACAGCTGTATTAACGCTGGAAGATGCAGAAATGATCAAGA 660  
Db 793 GGTGTGCCCAAGATTCACACAGCTGTATTAACGCTGGAAGATGATGATGTCTGGA 852  
Qy 661 ATGGCTTCTCATGGGATCAAAATTTGTCATTCAGCTAAAGATGGGGGCGAAAGCTACCCA 720  
Db 853 ATGGCTTCTCTGGGAGCAAAATTTGTCATTCAGCTAAAGATGGGGGCGAAAGCTATATCA 912  
Qy 721 GATACGTATCTCTCAACACTGTAGCAGAGATCACTGGGAGCAAAATTCACAAAGCTT 780  
Db 913 GATACAGATCTCTTCAACACTGTGCAAGATCACTGGGAGCAAAATTCACAAAGAAATTT 972  
Qy 781 GTACTGTGCTGAGCATCTGAGACAGCTGGGAGTGTGGGCGAGGTGCCATGATGATGCC 840  
Db 973 GTCTGTGTGAGTGCATCTGAGACAGCTGGGAGTGTGGGCGAGGTGCCATGATGATGCC 1032  
Qy 841 GGTGAGCTTATATCATAGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGTGCTCA 900  
Db 1033 GGTGAGCTTATATCATAGGGAAGCACTCTCACTTATTAAGATCTTGGGCTGTGCTCA 1092  
Qy 901 AAGAGACTCTGCGGCTGTGCTCTGAGCTGCAAGAGAAACAAGTGAATGTGTGCTTC 960  
Db 1093 AAGAGACTCTGCGGCTGTGCTCTGAGCTGCAAGAGAAACAAGTGAATGTGTGCTTC 1152  
Qy 961 CAGTATTTACAGTTACACAAGGTAATTTTTCAACTACAGTCTGTGTATGAGAGTCTGAC 1020  
Db 1153 CAGTATTTACAGTTACACAAGGTAATTTTTCAACTACAGTCTGTGTATGAGAGTCTGAC 1212  
Qy 1021 GCAGGAACCTTCTTAACCACTGGGCTGCAATTCATGCAAGTGAAGAAAGGCGAGGCTATC 1080  
Db 1213 TCAGGAACCTTCTTAACCACTGGGCTGCAATTCATGCAAGTGAAGAAAGGCGAGGCTATC 1272  
Qy 1081 ATGAGAGGTTATAGGCTGTGAGGCTGCAATTCATCACTCAAGTCTGAGGCAATGGA 1140  
Db 1273 ATGAGAGGTTATAGGCTGTGAGGCTGCAATTCATCACTCAAGTCTGAGGCAATGGA 1332  
Qy 1141 GAAGGAGCAGACATCACTTTTGGATCCAGCTGAGAGTCTGGAAGCACTGTACTTAT 1200  
Db 1333 GAAGGAGCAGACATCACTTTTGGATCCAGCTGAGAGTCTGGAAGCACTGTACTTAT 1392  
Qy 1201 GACTTATCAGTATTTCTTCTGATACCTCCACGAGACACCATATGATGATGAT 1260  
Db 1393 GACTTATCAGTATTTCTTCTGATACCTCCACGAGACACCATATGATGATGAT 1452  
Qy 1261 CCAAGCAGATGATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGAC 1320

Db 1453 CAAAGCAGATGATGTGCTGCTGCTGTTGGGCTGTTCGCTTACGTGTGGCAGAC 1512  
Qy 1321 ATGAGAGAAATGCTGCTAGGTCC 1344  
Db 1513 ATGAGAGAAATGCTGCCAGGTCC 1536  
  
RESULT 4  
US-09-984-271-95  
; Sequence 95, Application US/09984271  
; Publication No. US20030040088A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 71 Human Secreted Proteins  
; FILE REFERENCE: P2030P1  
; CURRENT APPLICATION NUMBER: US/09/984,271  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 09/482,273  
; PRIOR FILING DATE: 2000-01-13  
; PRIOR APPLICATION NUMBER: PCT/US99/15849  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: 60/092,921  
; PRIOR FILING DATE: 1998-07-15  
; PRIOR APPLICATION NUMBER: 60/092,922  
; PRIOR FILING DATE: 1998-07-15  
; PRIOR APPLICATION NUMBER: 60/092,956  
; PRIOR FILING DATE: 1998-07-15  
; NUMBER OF SEQ ID NOS: 267  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 95  
; LENGTH: 1134  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-984-271-95  
  
Query Match 57.5%; Score 772.4; DB 11; Length 1134;  
Best Local Similarity 99.7%; Pred. No. 4,9e-241;  
Matches 784; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
  
Qy 559 TCCTATTCAGTCTCTCAACAGGATTTTCAGAAATACAGAGTGGCGTCCCAAAATTC 618  
Db 1 TCCTATTCAGTCTCTCTCAACAGGATTTTCAGAAATACAGAGTGGCGTCCCAAAATTC 60  
Qy 619 ACAGCTGTATTCAGTGGGAAGATGCAAGATGATCAAGATAGCTTCTCATGGATC 678  
Db 61 ACAGCTGTATTCAGTGGGAAGATGCAAGATGATGCAAGATAGCTTCTCATGGATC 120  
Qy 679 AAAATGTCTACGCTTAAAGATGGGGCAAGACTTACCAGATGATTCCTTCAAC 738  
Db 121 AAAATGTCTACGCTTAAAGATGGGGCAAGACTTACCAGATGATTCCTTCAAC 180  
Qy 739 ACTGTACAGAGATCACTGGGAGCAATATTCAGAAACAGGTTGTACTGTGTACGTGACAT 798  
Db 181 ACTGTACAGAGATCACTGGGAGCAATATTCAGAAACAGGTTGTACTGTGTACGTGACAT 240  
Qy 799 CTGAGACGCTGGAGTGTGGGCAAGGCTGCAATGATGAGTGGGCTTATATCA 858  
Db 241 CTGAGACGCTGGAGTGTGGGCAAGGCTGCAATGATGAGTGGGCTTATATCA 300  
Qy 859 TGGAGACCTCTCACTTATTAAGATCTTGGCTGCTCAAGAGACTCTGCGCTG 918  
Db 301 TGGAGACCTCTCACTTATTAAGATCTTGGCTGCTCAAGAGACTCTGCGCTG 360  
Qy 919 GTGCTGTGAGCTGCAAGAAACAAGGTGTGAGTGTGCTTCCAGTATTTACGTACAC 978  
Db 361 GTGCTGTGAGCTGCAAGAAACAAGGTGTGAGTGTGCTTCCAGTATTTACGTACAC 420  
Qy 979 AAGTAAATATTTTCACTACGCTGTGATGAGTCTGAGAGGAACTTCTTACC 1038  
Db 421 AAGTAAATATTTTCACTACGCTGTGATGAGTCTGAGAGGAACTTCTTACC 480  
Qy 1039 ACTGGCTGCAATTCATGTCAGTGAAGAGGCGCATTCATGAGAGTTATGAGC 1098



```
LENGTH: 134
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (3)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (9)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (29)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (114)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (118)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (126)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-12141
```

Query Match 6.8%; Score 91; DB 10; Length 134;

Best Local Similarity 83.5%; Pred. No. 5.6e-19; Matches 11; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

```
QY 463 AACTACTCAGAGCGGTGCAATATCCGAAAGCGAGGCGCGTGGAGAGCTGCCAAGGTGGGG 522
DB 2 ANCTACTNAGAGCGGTGCAATATCCGAAAGCGAGGCGCGTGGAGAGCTGCCAAGGTGGGG 61
QY 523 GCTTTGCGATCTCTCTATTCG-ATCCGTGGCTCTCTTCCATCTACAGTCTCACACAGG 581
DB 62 GCTTTGCGATCTCTTATTTAGGATCCGTGGCTCTCTTCCATCTAAGCTCTMAANAGG 121
QY 582 TATTCAGGATAC 594
DB 122 TTTTNNAGGATAC 134
```

## RESULT 8

```
US-09-908-975-25931
; Sequence 25931, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25931
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-25931
```

Query Match 4.5%; Score 60.2; DB 13; Length 65;  
Best Local Similarity 95.4%; Pred. No. 4.2e-09;

```
Matches 62; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1261 CCNAGCAGATGATGTTGCGCTGCTGTTGGCTGTTCTTATGTTGACAGC 1320
DB 1 CCNAGCAGATGATGTTGCGCTGCTGTTGAGGCTTGTGCTTATGTTGACAGC 60
QY 1321 ATGGA 1325
DB 61 ATGGA 65
```

## RESULT 9

```
US-09-908-975-8960
; Sequence 8960, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8960
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-8960
```

Query Match 4.5%; Score 60; DB 13; Length 60;

Best Local Similarity 100.0%; Pred. No. 4.6e-09; Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 80 TCAACCTAGCTGTTATGTAAGCCGAGAACAGATCTTATGAGCATTTGGACTTCTGG 139
DB 1 TCAACCTAGCTGTTATGTAAGCCGAGAACAGATCTTATGAGCATTTGGACTTCTGG 60
```

## RESULT 10

```
US-10-059-585-5
; Sequence 5, Application US/10059585
; Publication No. US20030082776A1
; GENERAL INFORMATION:
; APPLICANT: Ota, Toshio
; APPLICANT: Isogai, Takao
; APPLICANT: Nishikawa, Tetsuo
; APPLICANT: Hayashi, Koji
; APPLICANT: Otsuka, Kaoru
; APPLICANT: Yamamoto, Jun-ichi
; APPLICANT: Ishii, Shizuko
; APPLICANT: Sugiyama, Tomoyasu
; APPLICANT: Wakamatsu, Ai
; APPLICANT: Nagai, Keiichi
; APPLICANT: Otsuki, Tetsuji
; APPLICANT: Funahashi, Shin-ichi
; APPLICANT: Senoo, Chiaki
; APPLICANT: Nezu, Jun-ichi
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEIN
; FILE REFERENCE: 06501-098001
; CURRENT APPLICATION NUMBER: US/10/059,585
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/JP00/05060
; PRIOR FILING DATE: 2000-07-28
```



;; CURRENT APPLICATION NUMBER: US/10/232,484  
;; CURRENT FILING DATE: 2002-08-30  
;; PRIOR APPLICATION NUMBER: 09/695,783  
;; PRIOR FILING DATE: 2000-10-24  
;; PRIOR APPLICATION NUMBER: 09/552,929  
;; PRIOR FILING DATE: 2000-04-18  
;; NUMBER OF SEQ ID NOS: 16  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO: 5  
;; LENGTH: 1363  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (848)..(1090)  
;; OTHER INFORMATION:  
US-10-232-484-5

Query Match 2.9%; Score 39.4; DB 15; Length 1363;  
Best Local Similarity 49.3%; Pred. No. 0.17;  
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 144 TACTGTGGACCCAGACTGAGTGGCTCCAGAAGCCTAGAAAAAGCCATCCAAATTATGTA 203  
DB 877 TAGTGTCTGTTCCAAACGCACTGTGTCAAGAGCTAGAAATTAACATTAGAGGCACAAC 936  
QY 204 CCAAAACCTGCAGCAAGATGGGCTGGAGAAAGTTACCTGGAGCCAGTGAAGAAATCCCA 263  
DB 937 CCTGAGAAATACACAAGGGGGGACCTTCCAGTAAGTGTGGAGAGAGAGAGGCGAG 996  
QY 264 CTGGAGAGGGGAGAGAAATCAGCTGTGATCTGGAGCCAAAGATTCAATTAAGATAGCAT 323  
DB 997 AGGGAGACAGGGGACAGGATTCAGCTTTGTGTGGGCTTCAGAGGTTCTTCAACAGGGGTAG 1056  
QY 324 CCTGGGCTTGGCAGCAGCATTTGGGACTC 352  
DB 1057 CCAGGATCTGGGAAACAGATCAGCGACTC 1085

RESULT 14  
US-10-044-090-579  
;; Sequence 579, Application US/10044090  
;; Publication No. US20020137081A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Olga Bandman  
;; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION  
;; FILE REFERENCE: PA-0028 US  
;; CURRENT APPLICATION NUMBER: US/10/044,090  
;; CURRENT FILING DATE: 2002-01-09  
;; NUMBER OF SEQ ID NOS: 850  
;; SOFTWARE: PERL Program  
;; SEQ ID NO: 579  
;; LENGTH: 2319  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; OTHER INFORMATION: Incyte ID No. US20020137081A1 023244.1  
;; LOCATION: 169, 173  
;; OTHER INFORMATION: a, t, c, g, or other  
US-10-044-090-579

Query Match 2.9%; Score 39.4; DB 14; Length 2319;  
Best Local Similarity 49.3%; Pred. No. 0.24;  
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 144 TACTGTGGACCCAGACTGAGTGGCTCCAGAAGCCTAGAAAAAGCCATCCAAATTATGTA 203  
DB 1560 TAGTGTCTGTTCCAAACGCACTGTGTCAAGAGCTAGAAATTAACATTAGAGGCACAAC 1619  
QY 204 CCAAAACCTGCAGCAAGATGGGCTGGAGAAAGTTACCTGGAGCCAGTGAAGAAATCCCA 263

DB 1620 CCTGAGAAATACACAAGGGGCAAGCTTCCAGTAGATGTGTGGAGAGAGAGGCGCAG 1679  
QY 264 CTGGAGAGGGGAGAGAAATCAGCTGTGATCTGGAGCCAAAGATTCAATTAAGATAGCAT 323  
DB 1680 AGGGAGACAGGGGACAGGATTCAGCTTTGTGTGGGCTTCAGAGGTTCTTCAACAGGGGTAG 1739  
QY 324 CCTGGGCTTGGCAGCAGCATTTGGGACTC 352  
DB 1740 CCAGGATCTGGGAAACAGATCAGCGACTC 1768

RESULT 15  
US-09-814-353-20277  
;; Sequence 20277, Application US/09814353  
;; Publication No. US20030165831A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Lee, John  
;; APPLICANT: Thompson, Pamela  
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
;; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
;; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
;; FILE REFERENCE: MRI-0068  
;; CURRENT APPLICATION NUMBER: US/09/814,353  
;; CURRENT FILING DATE: 2001-03-21  
;; PRIOR APPLICATION NUMBER: US 60/191,031  
;; PRIOR FILING DATE: 2000-03-21  
;; PRIOR APPLICATION NUMBER: US 60/207,124  
;; PRIOR FILING DATE: 2000-05-25  
;; PRIOR APPLICATION NUMBER: US 60/211,940  
;; PRIOR FILING DATE: 2000-06-15  
;; PRIOR APPLICATION NUMBER: US 60/216,820  
;; PRIOR FILING DATE: 2000-07-07  
;; PRIOR APPLICATION NUMBER: US 60/220,661  
;; PRIOR FILING DATE: 2000-07-25  
;; PRIOR APPLICATION NUMBER: US 60/257,672  
;; PRIOR FILING DATE: 2000-12-21  
;; NUMBER OF SEQ ID NOS: 22037  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO: 20277  
;; LENGTH: 3322  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-814-353-20277

Query Match 2.9%; Score 39.4; DB 13; Length 3322;  
Best Local Similarity 49.3%; Pred. No. 0.3;  
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 144 TACTGTGGACCCAGACTGAGTGGCTCCAGAAGCCTAGAAAAAGCCATCCAAATTATGTA 203  
DB 1946 TAGTGTCTGTTCCAAACGCACTGTGTCAAGAGCTAGAAATTAACATTAGAGGCACAAC 2005  
QY 204 CCAAAACCTGCAGCAAGATGGGCTGGAGAAAGTTACCTGGAGCCAGTGAAGAAATCCCA 263  
DB 2006 CCTGAGAAATACACAAGGGGCAAGCTTCCAGTAGATGTGTGGAGAGAGAGGCGCAG 2065  
QY 264 CTGGAGAGGGGAGAGAAATCAGCTGTGATCTGGAGCCAAAGATTATTAAGATAGCAT 323  
DB 2066 AGGGAGACAGGGGACAGGATTCAGCTTTGTGTGGGCTTCAGAGGTTCTTCAACAGGGGTAG 2125  
QY 324 CCTGGGCTTGGCAGCAGCATTTGGGACTC 352  
DB 2126 CCAGGATCTGGGAAACAGATCAGCGACTC 2154

Search completed: December 22, 2003, 23:44:41  
Job time: 421.737 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:18:08 ; Search time 2046.81 Seconds

(without alignments)  
11332.658 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_57\_623

Perfect score: 567

Sequence: 1 CCTATCAGATTATCTTACCA.....AGATTGTTGTTATACCAA 567

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapept 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl1.\*

1: gb Da:\*

2: gb Htg:\*

3: gb In:\*

4: gb Om:\*

5: gb Ov:\*

6: gb Pat:\*

7: gb Ph:\*

8: gb Pl:\*

9: gb Pr:\*

10: gb Ro:\*

11: gb Sts:\*

12: gb Sy:\*

13: gb Un:\*

14: gb Vi:\*

15: em Ba:\*

16: em Fun:\*

17: em Hum:\*

18: em In:\*

19: em Mu:\*

20: em Om:\*

21: em Or:\*

22: em Ov:\*

23: em Pat:\*

24: em Ph:\*

25: em Pl:\*

26: em Ro:\*

27: em Sts:\*

28: em Un:\*

29: em Vi:\*

30: em Htg Hum:\*

31: em Htg Inv:\*

32: em Htg Other:\*

33: em Htg Mus:\*

34: em Htg Pln:\*

35: em Htg Rod:\*

36: em Htg Mam:\*

37: em Htg Vrt:\*

38: em Gy:\*

39: em Htgo Hum:\*

40: em Htgo Mus:\*

41: em Htgo Other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	567	100.0	1472	9 AF107834	AF107834 Homo sapi
2	567	100.0	1851	6 BD106411	BD106411 Secreted
3	567	100.0	1860	6 AX006440	AX006440 Sequence
4	567	100.0	1928	9 BC020689	BC020689 Homo sapi
5	563.8	99.4	1860	6 BD127520	BD127520 Primer fo
6	563.8	99.4	1860	9 AK075132	AK075132 Homo sapi
7	555	97.9	609	6 BD059610	BD059610 Secreted
8	549	96.8	1794	9 AF119386	AF119386 Homo sapi
9	526.2	92.8	1778	6 BD105816	BD105816 Secretory
10	512.2	90.3	642	6 BD125210	BD125210 Primer fo
11	512.2	90.3	642	6 BD126476	BD126476 Primer fo
12	472.4	83.3	156702	2 AC084223	AC084223 Homo sapi
13	472.4	83.3	157927	9 AP006278	AP006278 Homo sapi
14	472.4	83.3	171301	9 AC010859	AC010859 Homo sapi
15	400	70.5	1726	10 AF131077	AF131077 Rattus no
16	399.6	70.5	1778	6 AX400829	AX400829 Sequence
17	399.6	70.5	1778	10 AF097723	AF097723 Rattus no
18	395.6	69.8	1716	10 AF009513	AF009513 Mus muscu
19	395.6	69.8	1796	10 BC037067	BC037067 Mus muscu
20	394	69.5	1596	10 AF107835	AF107835 Mus muscu
21	327.4	57.7	167883	2 AC121026	AC121026 Rattus no
22	327.4	57.7	218874	2 AC110103	AC110103 Rattus no
23	323.6	57.1	214287	2 AC129951	AC129951 Mus muscu
24	296	52.2	423	9 AF107833	AF107833 Homo sapi
25	290.8	51.3	314	6 BD058492	BD058492 Secreted
26	284.6	50.2	462	6 BD105811	BD105811 Secretory
27	284.6	50.2	462	6 E62849	E62849 Secretary P
28	254.6	44.9	317	6 BD076934	BD076934 5' EST of
29	121	21.3	217	6 BD076930	BD076930 5' EST of
30	49.2	8.7	7218	6 I66494	I66494 Sequence 14
31	46.6	8.2	1729	3 AF077194	AF077194 Acanthoch
32	45.2	8.0	161084	2 AC024030	AC024030 Homo sapi
33	45.2	8.0	161488	9 AP003083	AP003083 Homo sapi
34	45.2	8.0	173003	2 AC024245	AC024245 Homo sapi
35	41.8	7.4	259599	2 AC098081	AC098081 Rattus no
36	41.8	7.4	262687	2 AC128747	AC128747 Rattus no
37	41.6	7.3	230426	9 AC017099	AC017099 Homo sapi
38	41.2	7.3	266459	2 AC099884	AC099884 Mus muscu
39	41	7.2	3909	1 BA011045	BA011045 Buchnera ap
40	41	7.2	11249	1 AB014105	AB014105 Shewanella
41	40.6	7.2	10792	1 AB015790	AB015790 Shewanella
42	40.6	7.2	183155	2 AC128618	AC128618 Rattus no
43	40.6	7.2	253140	2 AC094055	AC094055 Rattus no
44	40.2	7.1	110000	2 FPMAL6P1_05	Continuation (6 of
45	40.2	7.1	162065	2 AC117061	AC117061 Rattus no

# ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
AF107834	AF107834	Homo sapiens clone LCH1 aminopeptidase mRNA, complete cds.	AF107834	AF107834.1	GI:5442029		Homo sapiens (human)	1472 bp mRNA linear PRI 12-JUL-1999			
							Homo sapiens				
							Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
								1 (bases 1 to 1472)			
								Liu,C.H., Lan,B.Y. and Chang,L.Y.			
								Cloning of the human aminopeptidase gene			
								Unpublished			

REFERENCE 2 (bases 1 to 1472)  
AUTHORS Liu, C. H., Lin, B. Y. and Chang, L. Y.  
TITLE Direct Submission  
JOURNAL Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia  
Sinica, RM 433, 128, Yen-Chun-Yuan Road SEC 2, Taipei 11529, Taiwan  
FEATURES  
source  
1. 1472  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="Uchi"  
/tissue\_type="liver"  
/dev\_stage="fetus"  
54. 1472  
/note="secretory protein"  
/codon\_start=1  
/product="aminopeptidase"  
/protein\_id="A04321.1"  
/db\_xref="GI:5442030"  
/translation="MKFLIFAFPGVHLISCSGKAIKNGISKRTPEIKETIASGCS  
DVAKIINLAVYGAONSVERIALNDVTVPKSGSKLEKAIQIMQNIQDGLK  
VHAEPRAPHEHGBESAVMEPRHIAIIGLSSIGTPPEGILTAELVITFDELQ  
RRAEARKIVINQPIYNSRTVQYRTQGAVEAAKVALASLRVASFISYSPHTG  
IQEYQDGVPIPTACITVEDAEKMSRMSHGKIKIVIQKMGAKTVPDTSFNTVAEIT  
GSKYPEGVLVSGHLSMDVGCAMDGGAFAISWEALSLIKDGLRPRKTLRLVLT  
AEQGVGAFOYQOLHKVINSISLVMSDAGTLPGLPTGSGSKRAIMEEVMSLL  
QPLNITQVLSHGBGDIINPMIOAGVPGASLLDLIKYRFFHSHGDMTVDPRQMY  
AAAVAVSYVAAMEMLPRS"

BASE COUNT 411 a 325 c 381 g 355 t  
ORIGIN

Query Match 100.0%; Score 567; DB 9; Length 1472;  
Best Local Similarity 100.0%; Pred. No. 7,2e-147;  
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTATCAGATTATCTTAAAGAAAACCACTGGAAAAAATGAATTCCTTATCTTC 60  
DB 12 CCTATCAGATTATCTTAAAGAAAACCACTGGAAAAAATGAATTCCTTATCTTC 71  
QY 61 GCATTTTCGGTGTGTTACCTTTTATCCTGTGCTCTGGAAAAGCTATGCAAGAT 120  
DB 72 GCATTTTCGGTGTGTTACCTTTTATCCTGTGCTCTGGAAAAGCTATGCAAGAT 131  
QY 121 GGCACTCTTAAGAGACTTTTGAAGAAATAAAGAAATAGCAGCTGTGAGATGT 180  
DB 132 GGCACTCTTAAGAGACTTTTGAAGAAATAAAGAAATAGCAGCTGTGAGATGT 191  
QY 181 GCTAAAGCAATCATCAACTGCTGTTTATGTTAAAGCCAGAAAGATCTTATAGCGA 240  
DB 192 GCTAAAGCAATCATCAACTGCTGTTTATGTTAAAGCCAGAAAGATCTTATAGCGA 251  
QY 241 TTGGCACTTCTGTGTGATCTGTGACCCGACGATGAGTGGCTCCAGAAAGCTTGA 300  
DB 252 TTGGCACTTCTGTGTGATCTGTGACCCGACGATGAGTGGCTCCAGAAAGCTTGA 311  
QY 301 GGCATCCAAATTATGTACCAAAACCTGCGAGCAAGTGGGCTGAGAAAGTTCACCTGGAG 360  
DB 312 GGCATCCAAATTATGTACCAAAACCTGCGAGCAAGTGGGCTGAGAAAGTTCACCTGGAG 371  
QY 361 CCACTGAGAAATACCCCACTGGGAGAGGGGAGAAAGATCAGCTGTGATGCTGAGACCAAGA 420  
DB 372 CCACTGAGAAATACCCCACTGGGAGAGGGGAGAAAGATCAGCTGTGATGCTGAGACCAAGA 431  
QY 421 ATTATTAAGATAGCATCTCTGGGCTTGTGCGACGAGCATTTGGGATCTCTCCAGAAAGCAT 480  
DB 432 ATTATTAAGATAGCATCTCTGGGCTTGTGCGACGAGCATTTGGGATCTCTCCAGAAAGCAT 491  
QY 481 ACAGCAGAAAGTTCTGTGTGATGACCTCTTGTGATGAATCTGAGAAAGGGCTCCAGAAAGCA 540  
DB 492 ACAGCAGAAAGTTCTGTGTGATGACCTCTTGTGATGAATCTGAGAAAGGGCTCCAGAAAGCA 551  
QY 541 AGAGGAGAAAGTTGTGTTTATTAACCA 567  
DB 541 AGAGGAGAAAGTTGTGTTTATTAACCA 567

DB 552 AGAGGAGAAAGTTGTGTTTATTAACCA 578

RESULT 2  
LOCUS BD106411 1851 bp DNA linear PAT 18-SEP-2002  
DEFINITION Secreted proteolins and polynucleotides encoding them.  
ACCESSION BD106411  
VERSION BD106411.1 GI:23201229  
KEYWORDS JP 2002503955-A/2.  
SOURCE  
ORGANISM Chlamydia sp.  
Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
REFERENCE 1 (bases 1 to 1851)  
AUTHORS Jacobs, K., McCoy, J. M., Lavallie, E. R., Racie, L. A., Merberg, D., Treacy, M., Spaulding, V., and Agostino, M. J.  
TITLE Secreted proteins and polynucleotides encoding them  
JOURNAL Patent: JP 2002503955-A/2 05-FEB-2002;  
GENETICS INSTITUTE INC  
COMMENT PN JP 2002503955-A/2  
PD 05-FEB-2002  
PF 20-MAR-1998 JP 1998545874  
PR 21-MAR-1997 US 08/822167, 19-MAR-1998 US 09/044466 PI  
KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI  
DAVID MERBERG,  
PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC  
C12N15/12, C07K14/47, A61K38/17  
CC Strandedness: Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.

FEATURES  
source  
1. 1851  
/organism="Chlamydia sp."  
/mol\_type="genomic DNA"  
/db\_xref="taxon:35827"

BASE COUNT 531 a 433 c 438 g 469 t  
ORIGIN

Query Match 100.0%; Score 567; DB 6; Length 1851;  
Best Local Similarity 100.0%; Pred. No. 7,3e-147;  
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTATCAGATTATCTTAAAGAAAACCACTGGAAAAAATGAATTCCTTATCTTC 60  
DB 57 CCTATCAGATTATCTTAAAGAAAACCACTGGAAAAAATGAATTCCTTATCTTC 116  
QY 61 GCATTTTCGGTGTGTTACCTTTTATCCTGTGCTCTGGAAAAGCTATGCAAGAT 120  
DB 117 GCATTTTCGGTGTGTTACCTTTTATCCTGTGCTCTGGAAAAGCTATGCAAGAT 176  
QY 121 GGCACTCTTAAGAGACTTTTGAAGAAATAAAGAAATAGCAGCTGTGAGATGT 180  
DB 177 GGCACTCTTAAGAGACTTTTGAAGAAATAAAGAAATAGCAGCTGTGAGATGT 236  
QY 181 GCTAAAGCAATCATCAACTGCTGTTTATGTTAAAGCCAGAAAGATCTTATAGCGA 240  
DB 237 GCTAAAGCAATCATCAACTGCTGTTTATGTTAAAGCCAGAAAGATCTTATAGCGA 296  
QY 241 TTGGCACTTCTGTGTGATCTGTGACCCGACGATGAGTGGCTCCAGAAAGCTTGA 300  
DB 297 TTGGCACTTCTGTGTGATCTGTGACCCGACGATGAGTGGCTCCAGAAAGCTTGA 356  
QY 301 GGCATCCAAATTATGTACCAAAACCTGCGAGCAAGTGGGCTGAGAAAGTTCACCTGGAG 360  
DB 357 GGCATCCAAATTATGTACCAAAACCTGCGAGCAAGTGGGCTGAGAAAGTTCACCTGGAG 416  
QY 361 CCACTGAGAAATACCCCACTGGGAGAGGGGAGAAAGATCAGCTGTGATGCTGAGACCAAGA 420  
DB 372 CCACTGAGAAATACCCCACTGGGAGAGGGGAGAAAGATCAGCTGTGATGCTGAGACCAAGA 476  
QY 421 ATTATTAAGATAGCATCTCTGGGCTTGTGCGACGAGCATTTGGGATCTCTCCAGAAAGCAT 480  
DB 477 ATTATTAAGATAGCATCTCTGGGCTTGTGCGACGAGCATTTGGGATCTCTCCAGAAAGCAT 536

QY	481	ACAGAGAAGTTCTGCTGATGACCTCTTTCGATGAACGACGAGAAGGAGCCCTCAGAGCA	540
Db	537	ACAGCAGAAGTTCTGCTGATGACCTCTTTCGATGAACGACGAGAAGGAGCCCTCAGAGCA	596
QY	541	AGAGGGAAGATTGTTGTTATATACCA	567
Db	597	AGAGGGAAGATTGTTGTTATATACCA	623
RESULT 3			
LOCUS	AX006440		
DEFINITION	Sequence 5 from Patent WO0004157.		
ACCESSION	AX006440		
VERSION	AX006440.1	GI:9994575	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
JOURNAL	Neefs,J.M., Peeters,D.C. and Pangalos,M.		
FEATURES	Cloning and characterisation of novel mammalian peptidases		
Source	Patent: WO 0004157-A 5 27-JAN-2000;		
	JANSSEN PHARMACEUTICA NV (BE); NEEFS JEAN MARC EDMOND FERNAND (BE)		
	PEETERS DANIELE CELINE GEORGE (BE); PANGALOS MENLAS (GB)		
BASE COUNT	518 a 420 c 451 g 471 t		
ORIGIN			
Query Match	100.0%; Score 567; DB 6; Length 1860;		
Best Local Similarity	100.0%; Pred. No. 7.3e-147;		
Matches	567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	CCTATCAGATATCTTAAACAAGAAACCACTGGAAGAAAAATGAATTCCTTATCTTC	60
Db	83	CCATACAGATATCTTAAACAAGAAACCACTGGAAGAAAAATGAATTCCTTATCTTC	142
QY	61	GCATTTTTCGGTGGTTCACCTTTTATCCCTGTGCTCTGGGAAGCTATATGCAAGAT	120
Db	143	GCATTTTTCGGTGGTTCACCTTTTATCCCTGTGCTCTGGGAAGCTATATGCAAGAT	202
QY	121	GGCATCTCTAAGAGCACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT	180
Db	203	GGCATCTCTAAGAGCACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGTT	262
QY	181	GCTAAGCAATCATCAACTAGCTGTTTATGTTAAGCCAGAAACAGATCCTATAGCGA	240
Db	263	GCTAAGCAATCATCAACTAGCTGTTTATGTTAAGCCAGAAACAGATCCTATAGCGA	322
QY	241	TTGGCACTTCGTGATAGTATCTTTGACCCAGACTGAGGGCTCCAGAACCTTAGAAAA	300
Db	323	TTGGCACTTCGTGATAGTATCTTTGACCCAGACTGAGGGCTCCAGAACCTTAGAAAA	382
QY	301	GCCATCCAAATTATGTATCAAAAACCTGACAGCAAGATGGCTGAGAAAAGTTCACTTGAG	360
Db	383	GCCATCCAAATTATGTATCAAAAACCTGACAGCAAGATGGCTGAGAAAAGTTCACTTGAG	442
QY	361	CCAGTGAAGATTAACCCACATGGGAGAGGGGAGAGAAATACAGCTGTATGCTGTGAGGCCA	420
Db	443	CCAGTGAAGATTAACCCACATGGGAGAGGGGAGAGAAATACAGCTGTATGCTGTGAGGCCA	502
QY	421	ATTCTAAGATAGCCATCTGGGTTTGGCAGACAGCATTTGGAGACTCTCCAGAAAGCAAT	480
Db	503	ATTCTAAGATAGCCATCTGGGTTTGGCAGACAGCATTTGGAGACTCTCCAGAAAGCAAT	562
QY	481	ACAGAGAAGTTCTGCTGATGACCTCTTTCGATGAACGACGAGAAGGAGCCCTCAGAGCA	540

Db	Query Match	Score	DB 9	Length	1928
Db	563	ACACGAGAAAGTTCTG	9878	567	ACACGAGAAAGTTCTGTCATGACCTCTTTCATGACGAGAGAGGGCCTCAGAGCA 622
Db	541	AGAGGAAAGATGTTGTTT	9878	567	AGAGGAAAGATGTTGTTTATACCA 567
Db	623	AGAGGAAAGATGTTGTTT	9878	567	AGAGGAAAGATGTTGTTTATACCA 649
RESULT 4					
LOCUS	BC020689	1928 bp	mRNA	linear	PRI 22-JAN-2002
DEFINITION	Homo sapiens, plasma glutamate carboxypeptidase, clone MGC:22418				
ACCESSION	IMMG:4251802, mRNA, complete cds.				
KEYWORDS	BC020689				
SOURCE	BC020689.1 GI:18088383				
ORGANISM	MGC.				
REFERENCE	Homo sapiens (human)				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 1928)				
JOURNAL	Strausberg, R.				
REMARK	Direct Submission				
COMMENT	Submitted (03-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>				
	Contact: MGC help desk				
	Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a>				
	Tissue Procurement: ATCC/DC/DTP				
	cDNA Library Preparation: CLONTECH Laboratories, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)				
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305				
	Web site: <a href="http://www-shgc.stanford.edu">http://www-shgc.stanford.edu</a>				
	Contact: (Dickson, Mark) <a href="mailto:mcd@axll.stanford.edu">mcd@axll.stanford.edu</a>				
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.				
FEATURES	source				
	1..1928				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="LocustID:10404"				
	/db_xref="taxon:9606"				
	/clone="MGC:22418 IMMG:4251802"				
	/tissue_type="Skin, melanocytic melanoma, high MDR."				
	/clone_id="NH_MGC_62"				
	/lab_host="DH10B"				
	/note="vector: pDNR-LIB"				
	167..1585				
	/codon_start=1				
	/product="plasma glutamate carboxypeptidase"				
	/protein_id="AAH20689.1"				
	/db_xref="GI:18088384"				
	/translation="MKEFLFAFPGVHLLSLCSGKAKCKNGISKRTFEETKEIASCGG				
	VDVAAILNLAVYGEKAQNRSEYRLALVDVTGPPRISGSKNLEKAIQMYQNLQDGLER				
	VHLEPRVILPMYERGESAVMLPEIRHKLILGLSSISGPEGITAEVLVTSDELELAVL				
	RRASERGKIVVNPQYIVTSRTQVQGVAKGALASLIRSAVSNTISYBPHGG				
	IQEVDGQVPIPTKCTIVYEDAEKMSRASHGIRKIVLQKGAATYPTDSENTYAEIT				
	GSRTPEQVVLVSGHLSMDVQGGAMDGGAFISWEALSLIKDGLRPKRTLRVLWTF				
	AEKGQAFQYQYQHLARKVNI SNLVSNSDAGTLPGLQFTGSEKARALMEETWSLL				
	QPLNITGVLSHGEGTIDNFWIQAQGVPAASLIDDLKYKFPHSHGDTMTVMDPKQNVV				
	AAAVAVAVSYVAAMEMLPRS"				
BASE COUNT	554 a 435 c 466 g 473 t				
ORIGIN					

Best Local Similarity 100.0%; Pred. No. 7.3e-147;  
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTATCAGATTATCTTAACAGAAAACCACTGAGAAAAAATGAATTCCTTATCTTC 60  
DB 125 CCTATCAGATTATCTTAACAGAAAACCACTGAGAAAAAATGAATTCCTTATCTTC 184  
QY 61 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGAGAACTATATGCAAGAT 120  
DB 185 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGAGAACTATATGCAAGAT 244  
QY 121 GGCATCTCTAAGAGACCTTTTGAAGAAATAAAGAAATATGCAAGTGTGAGATGTT 180  
DB 245 GGCATCTCTAAGAGACCTTTTGAAGAAATAAAGAAATATGCAAGTGTGAGATGTT 304  
QY 181 GCTAAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCCAAGAAAGATCCATAGAGCA 240  
DB 305 GCTAAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCCAAGAAAGATCCATAGAGCA 364  
QY 241 TTGGCACTCTGCTGTGATACCTGTGGAACCCAGACTGAGTGGCTCCAGAAACCTAGAAAAA 300  
DB 365 TTGGCACTCTGCTGTGATACCTGTGGAACCCAGACTGAGTGGCTCCAGAAACCTAGAAAAA 424  
QY 301 GGCATCCAAATTATGTATGCCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGGAG 360  
DB 425 GGCATCCAAATTATGTATGCCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGGAG 484  
QY 361 CCAAGTGAAGATACCCCACTGGAGAGGGAGAAAGATCACTGTGTATGCTGAGACCAAGA 420  
DB 485 CCAAGTGAAGATACCCCACTGGAGAGGGAGAAAGATCACTGTGTATGCTGAGACCAAGA 544  
QY 421 ATTCAATGAATATGACCTCTGGCTCTGGCAGCAGATGGAGCTCTCCAGAAAGCAATT 480  
DB 545 ATTCAATGAATATGACCTCTGGCTCTGGCAGCAGATGGAGCTCTCCAGAAAGCAATT 604  
QY 481 ACAGCAGAAAGTTCTGTGTGTGACCTCTTTCGATGAACTGACAGAAAGGGCTCAGAAAGA 540  
DB 605 ACAGCAGAAAGTTCTGTGTGTGACCTCTTTCGATGAACTGACAGAAAGGGCTCAGAAAGA 664  
QY 541 AGAGGGAAGATTGTTGTTTATTAACCA 567  
DB 665 AGAGGGAAGATTGTTGTTTATTAACCA 691

## RESULT 5

LOCUS BD127520 1860 bp DNA linear PAT 18-SEP-2002

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD127520.1 GI:23222465

VERSION JP 2002017375-A/2951

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Primer for synthesizing full-length cDNA and use thereof  
Patent: JP 2002017375-A 2951 22-JAN-2002;  
HEBIX RESEARCH INSTITUTE  
OS Homo sapiens (human)  
PN JP 2002017375-A/2951  
PD 22-JAN-2002  
PF 07-JUL-2000 JP 2000253172  
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO  
PI ISHII,  
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI  
SHINICHI KOUJIMA,  
PI TETSUJI OTSUKI, HISASHI KOGA  
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/PC

10,  
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof PH Key

FEATURES  
source  
1.1860  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT  
515 a 423 c 451 g 471 t  
ORIGIN

Query Match 99.4%; Score 563.8; DB 6; Length 1860;  
Best Local Similarity 99.6%; Pred. No. 5.7e-146;  
Matches 565; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCTATCAGATTATCTTAACAGAAAACCACTGAGAAAAAATGAATTCCTTATCTTC 60  
DB 87 CCTATCAGATTATCTTAACAGAAAACCACTGAGAAAAAATGAATTCCTTATCTTC 146  
QY 61 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGAGAACTATATGCAAGAT 120  
DB 147 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGAGAACTATATGCAAGAT 206  
QY 121 GGCATCTCTAAGAGACCTTTTGAAGAAATAAAGAAATATGCAAGTGTGAGATGTT 180  
DB 207 GGCATCTCTAAGAGACCTTTTGAAGAAATAAAGAAATATGCAAGTGTGAGATGTT 266  
QY 181 GCTAAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCCAAGAAAGATCCATAGAGCA 240  
DB 267 GCTAAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCCAAGAAAGATCCATAGAGCA 326  
QY 241 TTGGCACTCTGCTGTGATACCTGTGGAACCCAGACTGAGTGGCTCCAGAAAGCAATT 300  
DB 327 TTGGCACTCTGCTGTGATACCTGTGGAACCCAGACTGAGTGGCTCCAGAAAGCAATT 386  
QY 301 GGCATCCAAATTATGTATGCCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGGAG 360  
DB 387 GGCATCCAAATTATGTATGCCAAAACCTGCAGCAAGATGGGCTGAGAAAGTTCACTGGAG 446  
QY 361 CCAAGTGAAGATACCCCACTGGAGAGGGAGAAAGATCACTGTGTATGCTGAGACCAAGA 420  
DB 447 CCAAGTGAAGATACCCCACTGGAGAGGGAGAAAGATCACTGTGTATGCTGAGACCAAGA 506  
QY 421 ATTCAATGAATATGACCTCTGGCTCTGGCAGCAGATGGAGCTCTCCAGAAAGCAATT 480  
DB 507 ATTCAATGAATATGACCTCTGGCTCTGGCAGCAGATGGAGCTCTCCAGAAAGCAATT 566  
QY 481 ACAGCAGAAAGTTCTGTGTGTGACCTCTTTCGATGAACTGACAGAAAGGGCTCAGAAAGA 540  
DB 567 ACAGCAGAAAGTTCTGTGTGTGACCTCTTTCGATGAACTGACAGAAAGGGCTCAGAAAGA 626  
QY 541 AGAGGGAAGATTGTTGTTTATTAACCA 567  
DB 627 AGAGGGAAGATTGTTGTTTATTAACCA 653

## RESULT 6

LOCUS AK075132

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AK075132 1860 bp mRNA linear PRI 03-SEP-2002  
Homo sapiens cDNA FLJ90651 fis, clone PLAC1004482, moderately  
similar to Rattus norvegicus hematopoietic lineage switch 2 related  
protein (H182-rp) mRNA.  
AK075132  
AK075132.1 GI:22761022  
oligo capping; f1s (full insert sequence).  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,

Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,  
Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,  
Kojima, S., Nagahari, K., Masuno, Y., Ono, T., Okano, K., Yoshikawa, Y.,  
Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and  
Ninomiya, K.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 1860)  
Isogai, T. and Otsuki, T.  
Direct Submission  
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,  
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing;  
Research Association for Biotechnology; cDNA library construction;  
Institute of Medical Science, University of Tokyo, Laboratory of  
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass  
sequencing and clone selection; Helix Research Institute (supported  
by Japan Key Technology Center etc.).  
Location/Qualifiers  
1..1860  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLACE1004482"  
/issue\_type="placenta"  
/clone\_lib="PLACE1"  
/note="Cloning vector: PME18SFL3"  
129..1547  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="PAC1423.1"  
/db\_xref="GI:22761023"  
/translation="MKFLIFAFGGLSLICSGKAI CKNGISRTFEEIEEETASCG  
DVAKAIINLVYGAQNSYERLALVDVTPRSGSKNEKAQIMYQNIQDGLER  
VHLEPVRI PHMERGESAVMLEPRILKAIIGLSSIGTEPEGITAVLVTSFDEIK  
RRASEARGKIVVNOPIYNSRTVYRQGVAAKVALASILRSVAFSPHNG  
IOEYODGPKIPKACITVEPAEMSRMSHGKIKVIOIKGAKYPTDSEPNYAEIT  
GSKTPEOVVLSHLSMDVQGMADGGGAFISWLSLIKIGLPKRTLRVLT  
AEQGGVAFQYGLHKNVINSYLSVMSDAGTFLPTGLPTGSEKRAIVVEWMSLL  
QPLNITGVLSHGEIDINFWIOAGVPASLIDLKYKFFPHSHGDTMTWDPKQNV  
AAAVWAVSYVADMEEMLPKRS"

BASE COUNT 515 a 423 c 451 g 471 t  
ORIGIN

Query Match 99.4%; Score 563.8; DB 9; Length 1860;  
Best Local Similarity 99.6%; Pred. No. 5.7e-146;  
Matches 565; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCTATCAGATTATCTTAAGAAAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 60  
DB 87 CCTATCAGATTATCTTAAGAAAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 146  
QY 61 GCATTTTGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120  
DB 147 GCATTTTGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 206  
QY 121 GGCATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTT 180  
DB 207 GGCATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTT 266  
QY 181 GCTAAAGCAATCATCACTTACCTGTTTATATGTTAAAGCCCAAGACATCTTATGAGCGA 240  
DB 267 GCTAAAGCAATCATCACTTACCTGTTTATATGTTAAAGCCCAAGACATCTTATGAGCGA 326  
QY 241 TTGGCACTTCTGTTGATGTTGAGCCAGACTGAGTGGCTCCAAAGACCTTGAAGAAA 300  
DB 327 TTGGCACTTCTGTTGATGTTGAGCCAGACTGAGTGGCTCCAAAGACCTTGAAGAAA 386  
QY 301 GCCATCCAAATTAATATACCAAACTCGAGCAAGATGGGCTGAGAAAGTTTCACTGGAG 360  
DB 387 GCCATCCAAATTAATATACCAAACTCGAGCAAGATGGGCTGAGAAAGTTTCACTGGAG 446

QY 361 CCAGTAGAATATCCCACTGTGGAGAGGAGAAAGATCAGCTGTGATGTGAGCCCAAG 420  
DB 447 CCAGTAGAATATCCCACTGTGGAGAGGAGAAAGATCAGCTGTGATGTGAGCCCAAG 506  
QY 421 ATTCAATAGATAGCCATCTCTGCTTGGAGAGCAGATTTGGAGCTCTCCAGAGCAATT 480  
DB 507 ATTCAATAGATAGCCATCTCTGCTTGGAGAGCAGATTTGGAGCTCTCCAGAGCAATT 566  
QY 481 ACAGCAGAACTTCTGTTGTTGACCTTTTGGATGAATCTGAGAGAGGCTCAGAGCA 540  
DB 567 ACAGCAGAACTTCTGTTGTTGACCTTTTGGATGAATCTGAGAGAGGCTCAGAGCA 626  
QY 541 AGAGGAAATGTTGTTTATTAACCA 567  
DB 627 AGAGGAAATGTTGTTTATTAACCA 653

RESULT 7  
BD059610  
LOCUS 609 bp DNA linear PAT 27-AUG-2002  
DEFINITION Secreted expressed sequence tags (ESTs).  
ACCESSION BD059610  
VERSION BD059610.1 GI:22605216  
KEYWORDS JP 2001519666-A/1465.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 609)  
JACOBS, K., MCCOY, J.M., LAVALLIE, E.R., RACIE, L.A., MERBERG, D.,  
TREACY, M., SPAULDING, V. and AGOSTINO, M.J.  
Secreted expressed sequence tags (ESTs)  
Patent: JP 2001519666-A 1465 23-OCT-2001;  
GENETICS INSTITUTE INC  
PN JP 2001519666-A/1465  
PD 23-OCT-2001  
PR 10-APR-1998 JP 1998543068  
PR 10-APR-1997 US 08/835913  
PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI  
DAVID MERBERG.  
PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC  
C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:  
Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.  
FEATURES  
source 1..609  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"  
BASE COUNT 190 a 127 c 148 g 144 t  
ORIGIN

Query Match 97.9%; Score 555; DB 6; Length 609;  
Best Local Similarity 99.8%; Pred. No. 1.5e-143;  
Matches 566; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CCTATCAGATTATCTTAAGAAAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 60  
DB 29 CCTATCAGATTATCTTAAGAAAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 88  
QY 61 GCATTTTGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120  
DB 89 GCATTTTGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 148  
QY 121 GGCATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTT 180  
DB 149 GGCATCTCTAAGAGGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTT 208  
QY 181 GCTAAAGCAATCATCACTTACCTGTTTATATGTTAAAGCCCAAGACATCTTATGAGCGA 240

Db 209 GCTAAGCAATCATCAACCTAGCTGTTATGTGTAAAGCCCAAGACATCTATGAGCA 268  
 QY 241 TTGGACCTTCGTGTGATATCTGTGTGACCCAGAGTGGCTCCAGAACTAGAAAA 300  
 Db 269 TTGGACCTTCGTGTGATATCTGTGTGACCCAGAGTGGCTCCAGAACTAGAAAA 327  
 QY 301 GCCATCCAAATTTATGTACCAAAACCTGCAGCAAGATGGCTGGAGAAATTCCTGAG 360  
 Db 328 GCCATCCAAATTTATGTACCAAAACCTGCAGCAAGATGGCTGGAGAAATTCCTGAG 387  
 QY 361 CCAGTGAGATATCCCACTGGAGAGGGGAGAAATTCAGCTGTGATGCTGGAGCCAGA 420  
 Db 388 CCAGTGAGATATCCCACTGGAGAGGGGAGAAATTCAGCTGTGATGCTGGAGCCAGA 447  
 QY 421 ATTCATTAAGATAGCATCTGGGTCTTGGACAGACATGGGACTCTCCAGAGGCATT 480  
 Db 448 ATTCATTAAGATAGCATCTGGGTCTTGGACAGACATGGGACTCTCCAGAGGCATT 507  
 QY 481 ACAGCAGAAATTCCTGGTGTGACCTCTTTCAGATGACCTGCAGAGAGGCTCAGAGCA 540  
 Db 508 ACAGCAGAAATTCCTGGTGTGACCTCTTTCAGATGACCTGCAGAGAGGCTCAGAGCA 567  
 QY 541 AGAGGAGAAATTCCTGGTGTGACCTCTTTCAGATGACCTGCAGAGAGGCTCAGAGCA 567  
 Db 568 AGAGGAGAAATTCCTGGTGTGACCTCTTTCAGATGACCTGCAGAGAGGCTCAGAGCA 594  
 RESULT 8  
 AF119386 1794 bp mRNA linear PRI 20-MAY-1999  
 LOCUS AF119386 Homo sapiens blood plasma glutamate carboxypeptidase precursor  
 DEFINITION (PGCP) mRNA, complete cds.  
 ACCESSION AF119386  
 VERSION AF119386.1 GI:4877697  
 KEYWORDS  
 ORGANISM Homo sapiens (human)  
 SOURCE Homo sapiens  
 Bacteria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 1794)  
 Gindraas, R., Richard, C., El-Alfy, M., Morales, C.R., Potier, M. and Penezhetsky, A.V.  
 Purification, cDNA cloning, and expression of a new human blood plasma glutamate carboxypeptidase homologous to N-acetyl-aspartyl-alpha-glutamate carboxypeptidase/prostate-specific membrane antigen  
 J. Biol. Chem. 274 (17), 11742-11750 (1999)  
 MEDLINE 5923495  
 PUBMED 10206990  
 JOURNAL 2 (bases 1 to 1794)  
 Gindraas, R., Richard, C., El-Alfy, M., Morales, C.R., Potier, M. and Penezhetsky, A.V.  
 Direct Submission  
 Submitted (12-JAN-1999) Medical Genetics, Sainte-Justine Hospital, Montreal University, 3175 Cote Sainte-Catherine, Montreal, QU H3T 1C5, Canada  
 TITLE location/Qualifiers  
 JOURNAL 1..1794  
 ORGANISM="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /map="hg22.2; between DBS257 and DBS270"  
 1..1794  
 /gene="PGCP"  
 /gene="PGCP"  
 25..1650  
 /gene="PGCP"  
 /EC\_number="3.4.17.21"  
 /function="endopeptidase"  
 /note="member of M28 peptidase family; similar to glutamate carboxypeptidase II and to N-acetyl-aspartyl-alpha-glutamate carboxypeptidase (NALADase); similar to sequences with GenBank Accession

Numbers R18560 and G07404; prostate-specific membrane antigen (PSMA)"  
 /codon\_start=1  
 /product="blood plasma glutamate carboxypeptidase precursor"  
 /protein\_id="AA031418.1"  
 /db\_xref="GI:4877698"  
 /translation="MKFLIPAFPGVHLLSLGSKAIKNGISKRTPEEIKETIASCG DVAKATINILAVYGAQRNSYERIALVDVTGPRLSGKNLEKAIQIYVONIQDGLK VHEPVRIPIHREGEBSAAMLPEPIHKLITLIGSSITGEPIGTATVATVTSFDELQ RRASEARAKIVYNQPIYNTSRVQYRTQGAWEAKGALASLRVSASRISYSPHTG IOEYQDVSVPKIPNACITVEDAEWMSRASHGKIVLQKGAQYPTDSEFNVAET GSKYEQVAVVSGHLDSDMDGQAMDDGGAFLSWEALSLIKDLGPKKRLVLT AEQGVGAPQYQYQHLKVNISYSLVMSAGTFLPGLQFTGSEKARAIMESVSL QPLNITQVLSHGETDINFMVQAGVPASLLDDLYKYPFPHSHGDTMTWDPFRMLL ILFLGLFIMLTQWKKCCIGPRNSKETSCEPRLIGLQWMTPIHITISSSSSK HNSISCFLLSFLILSRKSRKNSHPLEPTT"  
 25..96  
 /gene="PGCP"  
 157..1647  
 /gene="PGCP"  
 /product="blood plasma glutamate carboxypeptidase"  
 205..207  
 /gene="PGCP"  
 /note="putative; glycosylation site"  
 559..561  
 /gene="PGCP"  
 /note="putative; glycosylation site"  
 1081..1083  
 /gene="PGCP"  
 /note="putative; glycosylation site"  
 1090..1092  
 /gene="PGCP"  
 /note="putative; glycosylation site"  
 1210..1212  
 /gene="PGCP"  
 /note="putative; glycosylation site"  
 BASE COUNT 522 a 391 c 422 g 457 t 2 others  
 ORIGIN  
 Query Match 96.8%; Score 549; DB 9; Length 1794;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-142;  
 Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 19 CAAGAAACCAACCTGGAAGAAATGAATTCCTTATCTTGGATTTTGGTGGT 78  
 Db 1 CAAGAAACCAACCTGGAAGAAATGAATTCCTTATCTTGGATTTTGGTGGT 60  
 QY 79 CACCTTTATCCCTGTGCTCTGGAAAGCTATATGCAAGATGCAATCTTAAGAGCA 138  
 Db 61 CACCTTTATCCCTGTGCTCTGGAAAGCTATATGCAAGATGCAATCTTAAGAGCA 120  
 QY 139 TTGAGAGAAATTAAGAAATATGCCAGCTGTGGAATGTGCTTAAGCAATATCAAC 198  
 Db 121 TTGAGAGAAATTAAGAAATATGCCAGCTGTGGAATGTGCTTAAGCAATATCAAC 180  
 QY 199 CTACCTGTTTATGTGTAAGCCAGAAAGCAAGATCCTATGAGGATGGCACTTGGTTGAT 258  
 Db 181 CTACCTGTTTATGTGTAAGCCAGAAAGCAAGATCCTATGAGGATGGCACTTGGTTGAT 240  
 QY 259 ACTGTGGAACCAAGATGAGTGGCTCCAGAAAGCAATTAAGCAATTAATATGATC 318  
 Db 241 ACTGTGGAACCAAGATGAGTGGCTCCAGAAAGCAATTAAGCAATTAATATGATC 300  
 QY 319 CAAAACCTGCAGCAAGATGGCTGGAGAAAGTCACTGAGCCAGTGAATATCCCCAC 378  
 Db 301 CAAAACCTGCAGCAAGATGGCTGGAGAAAGTCACTGAGCCAGTGAATATCCCCAC 360  
 QY 379 TGGAGAGGGAGAGAAAGATGCTGTGATGCTGAGAGCCAGAAATTCATTAAGATGCCATC 438  
 Db 361 TGGAGAGGGAGAGAAAGATGCTGTGATGCTGAGAGCCAGAAATTCATTAAGATGCCATC 420  
 QY 439 CTGGGCTTGGACACACATTTGGAGCTCTCCAGAAAGCATTAAGAGAAATTTCTGGTG 498

```

Db      421 CTGGGCTTGGCAGACATGAGACTCTCCAGAGGATTTACAGCAGAACTTCTGGTG 480
Qy      499 GTACACCTCTTCCATGATGACGACAGAGGGCTCCAGAGCAAGAGGAATGTTGTT 558
Db      481 GTGACCTCTTTCATGATGACGACAGAGGGCTCCAGAGCAAGAGGAATGTTGTT 540
Qy      559 TATACCAA 567
Db      541 TATACCAA 549

RESULT 9
LOCUS   BD105816
DEFINITION Secretory protein.
ACCESSION BD105816
VERSION   BD105816.1 GI:23200634
KEYWORDS  JP 2002502234-A/7.
SOURCE    synthetic construct
ORGANISM  artificial sequences.
REFERENCE 1 (bases 1 to 1778)
AUTHORS   Jacobs,K., McCoy,J.M., Racie,L.A., Lavallie,E.R., Werberg,D. and Spaulding,V.
TITLE     Secretory protein
JOURNAL   Patent: JP 2002502234-A 7 22-JAN-2002;
COMMENT   GENETICS INSTITUTE INC
          PN JP 2002502234-A/7
          PD 22-JAN-2002
          PF 16-APR-1997 JP 1997537384
          PR 18-APR-1996 US 08/634325 13-JAN-1997 US 08/783520 PI
          KENNETH JACOBS,JOHN M MCCOY,LISA A RACIE,EDWARD R LAVALLIE, PI
          DAVID WERBERG,
          PI VIKKI SPAULDING
          PC C12N15/12,C07K14/47,A61K38/17
          CC Strandedness: Double;
          CC Topology: Linear;
          FH Key Location/Qualifiers.

FEATURES
Source   1..1778
          /organism="synthetic construct"
          /mol type="genomic DNA"
          /db xref="taxon:32630"

BASE COUNT 514 a 386 c 422 g 456 t

Query Match 92.8%; Score 526.2; DB 6; Length 1778;
Best Local Similarity 98.5%; Pred. No. 1.7e-135;
Matches 531; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      29 AACTGGAAGAAAAATGAAATTCCTTATCTTGCGATTTTGGTGCTGTCACCTTTAT 88
Db      3 AGCTTGGACACAGAGTAAATTCCTTATCTTGCGATTTTGGTGCTGTCACCTTTAT 62
Qy      89 CCCGTGCTCGGGAAGAACTATATGCAAGATGCGATCTCAAGGAGCACTTTGAAGAA 148
Db      63 CCTGTGCTCGGGAAGAACTATATGCAAGATGCGATCTCAAGGAGCACTTTGAAGAA 122
Qy      149 TAAAGAAGAAATAGCCAGCTGTGAGATGTTGCTAAAGCAATCATCAACCTAGCTGTT 208
Db      123 TAAAGAAGAAATAGCCAGCTGTGAGATGTTGCTAAAGCAATCATCAACCTAGCTGTT 182
Qy      209 ATGTAAAGCCCAAGACAGATCTATGAGCGATTTGGCACTTCTGTTGATACTGTTGAC 268
Db      183 ATGTAAAGCCCAAGACAGATCTATGAGCGATTTGGCACTTCTGTTGATACTGTTGAC 242
Qy      269 CCAGACTAGTGGCTCCAGAAACCTAGAAAGAAAGCAATCAAAATTATGACAAAACCTGC 328
Db      243 CCAGACTAGTGGCTCCAGAAACCTAGAAAGAAAGCAATCAAAATTATGACAAAACCTGC 302
Qy      329 AGCAAGATGGGCTGAGAAAGTTCACTGAGCCAGTGAAGATACCCCACTGGAGAGGG 388

```

```

Db      303 AGCAAGATGGGCTGAGAAAGTTCACTGAGCCAGTGAAGATACCCCACTGGAGAGG 362
Qy      389 GAGAGAAATCAGCTGTATGCTGGAGCCAGAAATTCATTAAGATAGCATCTGGGCTTG 448
Db      363 GAGAGAAATCAGCTGTATGCTGGAGCCAGAAATTCATTAAGATAGCATCTGGGCTTG 422
Qy      449 GCAGCAGCAATTTGGACTCTCCAGAAAGCAATTAAGAGAAATTTGCTGGTGAACCTTT 508
Db      423 GCAGCAGCAATTTGGACTCTCCAGAAAGCAATTAAGAGAAATTTGCTGGTGAACCTTT 482
Qy      509 TCGATGAATCTGCAGAGAGGGCTCCAGAGCAAGAGGAAGATTGTTTATACCA 567
Db      483 TCGATGAATCTGCAGAGAGGGCTCCAGAGCAAGAGGAAGATTGTTTATACCA 541

RESULT 10
LOCUS   BD125210
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD125210
VERSION   BD125210.1 GI:23220155
KEYWORDS  JP 2002017375-A/641.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 642)
AUTHORS   Ota,T., Nishikawa,T., Isogai,T., Hayaishi,K., Ishii,S., Kawai,Y.,
          Makamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.
TITLE     Primer for synthesizing full-length cDNA and use thereof
JOURNAL   Patent: JP 2002017375-A 641 22-JAN-2002;
COMMENT   HELIX RESEARCH INSTITUTE
          OS Homo sapiens (human)
          PN JP 2002017375-A/641
          PD 22-JAN-2002
          PF 07-JUL-2000 JP 2000253172
          PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
          PI ISHII,
          PI YURI KAWAI,AI MAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
          SHINICHI KOJIMA,
          PI TETSUJI OTSUKI,HISASHI KOGA
          PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/00
          10',C12P21/02,C12P21/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
          Primer for synthesizing full-length cDNA and use thereof FH Key
          Location/Qualifiers
          FT source 1..642
          /organism="Homo sapiens (human)".
          /db xref="taxon:9606"

FEATURES
Source   1..642
          /organism="Homo sapiens"
          /mol type="genomic DNA"
          /db xref="taxon:9606"

BASE COUNT 183 a 144 c 171 g 139 t 5 others

Query Match 90.3%; Score 512.2; DB 6; Length 642;
Best Local Similarity 98.4%; Pred. No. 1.2e-131;
Matches 547; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

Qy      1 CCTATCAGATTAATCTTAAAGAAAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 60
Db      87 CCTATCAGATTAATCTTAAAGAAAGAAACCACTGGAAAAAATGAATTCCTTATCTTC 146
Qy      61 GCAATTTTCGGTGGTGTCACTTTTATCCCTGTGCTCTGGGAAGCAATATGCAAGAT 120
Db      147 GCAATTTTCGGTGGTGTCACTTTTATCCCTGTGCTCTGGGAAGCAATATGCAAGAT 206
Qy      121 GGCAATCTTAAGAGACTTTTGAAGAAATTAAGAAAGAAATAGCCAGCTGTGAGATGT 180
Db      207 GGCAATCTTAAGAGACTTTTGAAGAAATTAAGAAAGAAATAGCCAGCTGTGAGATGT 266

```

QY 181 GCTAAGCATCATCAACCTAGCTGTTTATGTTAAAGCCAGAACAGATCCTATGAGCGA 240  
 DB 267 GCTAAGCATCATCAACCTAGCTGTTTATGTTAAAGCCAGAACAGATCCTATGAGCGA 326  
 QY 241 TTGGCACTTCTGTTGATCTGTTGAGCCAGACTGATGCTCCAGAACCTTAAAGAAA 300  
 DB 327 TTGGCACTTCTGTTGATCTGTTGAGCCAGACTGATGCTCCAGAACCTTAAAGAAA 386  
 QY 301 GCCATCCAAATTTATGATCAAAAACCTGACGAGATGGGCTGAGAAAAGTTCACTGAG 360  
 DB 387 GCCATCCAAATTTATGATCAAAAACCTGACGAGATGGGCTGAGAAAAGTTCACTGAG 446  
 QY 361 CCAGTGAATATACCCCTGAGGAGAGGAGAGAAATGAGCTGATGCTGAGGAGCAAGA 420  
 DB 447 CCAGTGAATATACCCCTGAGGAGAGGAGAGAAATGAGCTGATGCTGAGGAGCAAGA 506  
 QY 421 ATTCAATTA-GATAGCATCTGCTGCTTGGCAGCAGATTTGGGACTCTCCAGAAAGGAT 479  
 DB 507 ATTCAATTAAGATAGCATCTGCTGCTTGGCAGCAGATTTGGGACTCTCCAGAAAGGAT 566  
 QY 480 TACAGCAGAAAGTTCTGTTGG-TGACCTCTTTGATGAAGTCTGACAGAA-GGGCTCAGAA 537  
 DB 567 TACAGCAGAAAGTTCTGTTGGGAGTGAACCTCTTTCATGAAGTCTGACAGAAAGGCTCAGAA 626  
 QY 538 GCAGAGGGAGAGATTG 553  
 DB 627 GCAGAGGGAGAGATTG 642

## RESULT 11

BD126476 642 bp DNA linear PAT 18-SEP-2002  
 LOCUS BD126476  
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
 ACCESSION BD126476  
 VERSION 1 GI:232221421  
 KEYWORDS JP 2002017375-A/1907.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 AUTHORS Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y., Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and Koga,H.  
 TITLE Primer for synthesizing full-length cDNA and use thereof  
 JOURNAL Patent: JP 2002017375-A 1907 22-JAN-2002.  
 COMMENT HELIX RESEARCH INSTITUTE  
 OS Homo sapiens (human)  
 PN JP 2002017375-A/1907  
 PD 22-JAN-2002 JP 2002025172  
 PF 07-JUL-2000 JP 2002025172  
 PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO PI  
 ISHII,  
 PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI  
 SHINICHI KOJIMA,  
 PI TETSUO OTSUKI, HISASHI KOGA

PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
 C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
 10.  
 PC C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
 Primer for synthesizing full-length cDNA and use thereof FH Key  
 FT source  
 FT 1. 642  
 Location/Qualifiers  
 /organism='Homo sapiens (human)'.  
 1. 642  
 /organism='Homo sapiens (human)'.  
 /mol\_type='genomic DNA'.  
 /db\_xref='taxon:9606'

BASE COUNT 183 a 144 c 171 g 139 t 5 others  
 ORIGIN

Query Match 90.3%; Score 512.2; DB 6; Length 642;  
 Best Local Similarity 98.4%; Pred. No. 1.2e-11;  
 Matches 547; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

QY 1 CTTATCAGATTATCTTAAAGAAACCACTGGAAAAAATGAAATTCCTTATCTTC 60  
 DB 87 CTTATCAGATTATCTTAAAGAAACCACTGGAAAAAATGAAATTCCTTATCTTC 146  
 QY 61 GCAATTTTGGGTGTGTTCACTTTTATCCCTGCTCTGGGAAAGCTATATGCAAGAT 120  
 DB 147 GCAATTTTGGGTGTGTTCACTTTTATCCCTGCTCTGGGAAAGCTATATGCAAGAT 206  
 QY 121 GGCATCTCTAAGAGACTTTTGAAGAAATTAAGAAATATGACCTGTGAGATGTT 180  
 DB 207 GGCATCTCTAAGAGACTTTTGAAGAAATTAAGAAATATGACCTGTGAGATGTT 266  
 QY 181 GCTAAGCATCATCAACCTAGCTGTTTATGTTAAAGCCAGAACAGATCCTATGAGCGA 240  
 DB 267 GCTAAGCATCATCAACCTAGCTGTTTATGTTAAAGCCAGAACAGATCCTATGAGCGA 326  
 QY 241 TTGGCACTTCTGTTGATCTGTTGAGCCAGACTGATGCTCCAGAACCTTAAAGAAA 300  
 DB 327 TTGGCACTTCTGTTGATCTGTTGAGCCAGACTGATGCTCCAGAACCTTAAAGAAA 386  
 QY 301 GCCATCCAAATTTATGATCAAAAACCTGACGAGATGGGCTGAGAAAAGTTCACTGAG 360  
 DB 387 GCCATCCAAATTTATGATCAAAAACCTGACGAGATGGGCTGAGAAAAGTTCACTGAG 446  
 QY 361 CCAGTGAATATACCCCTGAGGAGAGGAGAGAAATGAGCTGATGCTGAGGAGCAAGA 420  
 DB 447 CCAGTGAATATACCCCTGAGGAGAGGAGAGAAATGAGCTGATGCTGAGGAGCAAGA 506  
 QY 421 ATTCAATTA-GATAGCATCTGCTGCTTGGCAGCAGATTTGGGACTCTCCAGAAAGGAT 479  
 DB 507 ATTCAATTAAGATAGCATCTGCTGCTTGGCAGCAGATTTGGGACTCTCCAGAAAGGAT 566  
 QY 480 TACAGCAGAAAGTTCTGTTGG-TGACCTCTTTGATGAAGTCTGACAGAA-GGGCTCAGAA 537  
 DB 567 TACAGCAGAAAGTTCTGTTGGGAGTGAACCTCTTTCATGAAGTCTGACAGAAAGGCTCAGAA 626  
 QY 538 GCAGAGGGAGAGATTG 553  
 DB 627 GCAGAGGGAGAGATTG 642

## RESULT 12

AC084223/c 156702 bp DNA linear HTG 07-JUL-2001  
 LOCUS AC084223  
 DEFINITION Homo sapiens chromosome 8 clone RP11-80P10 map 8, WORKING DRAFT  
 SEQUENCE, 4 unordered pieces.  
 AC084223  
 AC084223 GI:14626321  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULFILLTOP.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 TITLE Homo sapiens chromosome 8, clone RP11-80P10  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 156702)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Barne,N., Baetien,V., Beda,F., Boguslavsky,L., Bonkagiller,B., Brown,A., Burkett,G., Campopiano,A., Caele,A., Chepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Iliev,I., Johnson,R., Jones,C., Kam,L., Karatas,A., Labocque,K., Lamazares,R., Landers,T., Lehoczy,D., Levine,R., Liu,C., Liu,G., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,



Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollard, V., Raymond, C., Rieback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schaner, S., Severy, P., Soguen, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Teffaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J., Vassiliev, H., Vei, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

## TITLE

Submitted (17-OCT-2000)

## JOURNAL

Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

## COMMENT

On Jul 7, 2001 this sequence version replaced gi:13249444.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: L11261

Center clone name: 80\_P\_10

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 154179 bases at least Q40

Consensus quality: 155367 bases at least Q30

Consensus quality: 155704 bases at least Q20

Insert size: 163000; agarose-fp

Insert size: 156402; sum-of-ctnigs

Quality coverage: 5.7 in Q20 bases; agarose-fp

Quality coverage: 6.0 in Q20 bases; sum-of-ctnigs

NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 12817: contig of 12817 bp in length  
\* 12818 12917: gap of 100 bp  
\* 12918 59770: contig of 46853 bp in length  
\* 59771 59870: gap of 100 bp  
\* 59871 113148: contig of 53278 bp in length  
\* 113149 113248: gap of 100 bp  
\* 113249 156702: contig of 43454 bp in length.

## FEATURES

## SOURCE

1. 156702

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="8"

/map="8"

/clone="RP11-80P10"

/clone\_lib="RPC1-11 Human Male BAC"

1. 12817

/note="assembly\_fragment"

clone\_end:SP6

vector\_side:left"

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

## ORIGIN

Query Match 83.3%; Score 472.4; DB 2; Length 156702;  
Best Local Similarity 99.84; Pred. No. 2,3e-120;  
Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

63583 AGATTATCTTACAGAGAAACCACTGGAAAAAATGAATTCCTTATCTTGCATTT 66  
|||||

67 TTGGTGTGTGTTACCTTTATTCCTGTCTGTGGAAAGCTATATGCAAGATGGCATC 126  
|||||

63523 TTGGTGTGTGTTACCTTTATTCCTGTCTGTGGAAAGCTATATGCAAGATGGCATC 63464  
|||||

127 TCTAGAGAGCTTTGAGAAATTAAGAAATATAGCCAGCTGTGAGATGTTGCTTAA 186  
|||||

63463 TCTAGAGAGCTTTGAGAAATTAAGAAATATAGCCAGCTGTGAGATGTTGCTTAA 63404  
|||||

187 GCATCATCAACCTAGCTGTTATGTGTAAGCCAGAACAGATCTATGAGCGATTGGCA 246  
|||||

63403 GCATCATCAACCTAGCTGTTATGTGTAAGCCAGAACAGATCTATGAGCGATTGGCA 63344  
|||||

247 CTTGTGTGTATCTGTTGAGACCCAGACTGAGTGTCTCAAGAACTTAAAGCAATC 306  
|||||

63343 CTTGTGTGTATCTGTTGAGACCCAGACTGAGTGTCTCAAGAACTTAAAGCAATC 63284  
|||||

307 CAATTATGTACCAAACTGCAGCAAGATGGCTGAGAAAGTTCACTGGAGCCAGTG 366  
|||||

63283 CAATTATGTACCAAACTGCAGCAAGATGGCTGAGAAAGTTCACTGGAGCCAGTG 63224  
|||||

367 AGATATCCCACTGGAGAGAGGAGAAATCACTGTGATGTGAGCCAGCAATTCAT 426  
|||||

63223 AGATATCCCACTGGAGAGAGGAGAAATCACTGTGATGTGAGCCAGCAATTCAT 63164  
|||||

427 AAGATGACCATCTGCTGCTTGTGAGAGCAATTTGGACTCTCTCAAGAGCATTT 480  
|||||

63163 AAGATGACCATCTGCTGCTTGTGAGAGCAATTTGGACTCTCTCAAGAGCATTT 63110  
|||||

RESULT 13  
AP006278 157927 bp DNA linear PRI 25-MAR-2003  
LOCUS Homo sapiens genomic DNA, chromosome 8q22.1, clone: KB1573B4,  
DEFINITION complete sequence.  
ACCESSION AP006278  
VERSION AP006278.1 GI:29170583  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 Shimizu, N. and Asakawa, S.  
Homo sapiens DNA chromosome 8 SEQUENCE  
JOURNAL Published Only in Database (2003)  
REFERENCE 2 (bases 1 to 157927)  
AUTHORS Shimizu, N. and Asakawa, S.  
TITLE Direct Submission  
JOURNAL Submitted (24-MAR-2003) Nobuyoshi Shimizu, Keio University, School  
of Medicine, Molecular Biology/ 35 Shinanomachi, Shinjuku-ku, Tokyo  
160-8582, Japan (E-mail: nshimizu@med.keio.ac.jp,  
Tel: 81-3-3351-2370, Fax: 81-3-3351-2370)  
FEATURES  
SOURCE  
1. 157927  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8q22.1"  
/clone="KB1573B4"  
/cell\_line="FLB 14 - 14"  
/cell\_type="pre-pro-B cell"  
/clone\_lib="Keio BAC library"

BASE COUNT 52002 a 30386 c 29018 g 44993 t 303 others

repeat_region	103. .202	/evidence=not experimental
repeat_region	/rpt_family="TlMe"	/rpt_family="TlA)n"
repeat_region	complement(217. .695)	/evidence=not experimental
repeat_region	/rpt_family="HERVL"	/rpt_family="At_rich"
repeat_region	complement(733. .1197)	/evidence=not experimental
repeat_region	/rpt_family="HERVL"	complement(117577. .17931)
repeat_region	complement(1135. .1381)	/evidence=not experimental
repeat_region	/rpt_family="HERVL"	/rpt_family="TlM4"
repeat_region	complement(1462. .1817)	/evidence=not experimental
repeat_region	/evidence=not experimental	/rpt_family="TlM4"
repeat_region	complement(1821. .1863)	/evidence=not experimental
repeat_region	/rpt_family="MLT2A1"	/rpt_family="TlM4"
repeat_region	1864. .1888	/evidence=not experimental
repeat_region	/rpt_family="TlM4"	/rpt_family="HER103"
repeat_region	1905. .1961	/evidence=not experimental
repeat_region	/rpt_family="TlA)n"	/rpt_family="MIR"
repeat_region	complement(1973. .2381)	complement(119949. .20245)
repeat_region	/evidence=not experimental	/evidence=not experimental
repeat_region	/rpt_family="MLT2A1"	/rpt_family="TlM4"
repeat_region	complement(2473. .3057)	/evidence=not experimental
repeat_region	/rpt_family="TlM4"	/evidence=not experimental
repeat_region	3064. .3092	/rpt_family="At_rich"
repeat_region	/evidence=not experimental	/rpt_family="TlM4"
repeat_region	/rpt_family="At_rich"	/evidence=not experimental
repeat_region	complement(3093. .9136)	complement(120530. .20781)
repeat_region	/evidence=not experimental	/evidence=not experimental
repeat_region	/rpt_family="TlPA3"	/rpt_family="HAL1"
repeat_region	complement(9154. .9483)	21107. .21135
repeat_region	/evidence=not experimental	/evidence=not experimental
repeat_region	/rpt_family="TlM4"	complement(121428. .21700)
repeat_region	9661. .9951	/evidence=not experimental
repeat_region	/evidence=not experimental	/rpt_family="HAL1"
repeat_region	/rpt_family="AluSg"	21701. .22034
repeat_region	9960. .10103	/evidence=not experimental
repeat_region	/evidence=not experimental	/rpt_family="AluSx"
repeat_region	/rpt_family="TlG"	complement(122035. .22140)
repeat_region	complement(10102. .10336)	/evidence=not experimental
repeat_region	/evidence=not experimental	/rpt_family="HAL1"
repeat_region	/rpt_family="MIR"	complement(124217. .24393)
repeat_region	10351. .10541	/evidence=not experimental
repeat_region	/evidence=not experimental	/rpt_family="MIR"
repeat_region	/rpt_family="TlG"	complement(124453. .24675)
repeat_region	11642. .11941	/evidence=not experimental
repeat_region	/evidence=not experimental	/rpt_family="MIR"
repeat_region	/rpt_family="AluSx"	complement(124890. .25101)
repeat_region	12296. .12351	/evidence=not experimental
repeat_region	/evidence=not experimental	/rpt_family="MIR"
repeat_region	/rpt_family="MIR3"	26204. .26409
repeat_region	12352. .12530	/evidence=not experimental
repeat_region	/evidence=not experimental	/rpt_family="MIR"
repeat_region	/rpt_family="MER5A"	26451. .26507
repeat_region	12531. .12575	/evidence=not experimental
repeat_region	/evidence=not experimental	/rpt_family="MIR"
repeat_region	/rpt_family="MIR3"	26979. .27110
repeat_region	13183. .13204	/evidence=not experimental
repeat_region	/evidence=not experimental	/rpt_family="MIR"
repeat_region	/rpt_family="At_rich"	complement(127349. .27525)
repeat_region	13487. .13604	/evidence=not experimental
repeat_region	/rpt_family="TlG"	/rpt_family="TlM4"
repeat_region	/evidence=not experimental	complement(127526. .27908)
repeat_region	14340. .14545	/evidence=not experimental
repeat_region	/evidence=not experimental	/rpt_family="MLT1B"
repeat_region	/rpt_family="TlMe"	complement(127909. .28143)
repeat_region	14753. .14795	/evidence=not experimental

```

repeat_region      /rpt_family="L1M4"
                    /evidence=not experimental
                    /rpt_family="L1M61"
repeat_region      29103..29231
                    /evidence=not experimental
                    /rpt_family="L1M10"
repeat_region      complement(30430..30598)
                    /evidence=not experimental
                    /rpt_family="L2"
repeat_region      30598..30833
                    /evidence=not experimental
                    /rpt_family="MIR"
repeat_region      complement(32270..32576)
                    /evidence=not experimental
                    /rpt_family="Alu5x"
repeat_region      complement(34260..34488)
                    /evidence=not experimental
                    /rpt_family="Alu4b"
repeat_region      35334..35382
                    /evidence=not experimental
                    /rpt_family="T1A"n"

```

Query Match 83.3%; Score 472.4; DB 9; Length 157927;  
 Best Local Similarity 99.8%; Pred. No. 2.3e-120;  
 Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 7 AGATTACTTAAACAAGAAACCACTGGAAGAAAGAAATGAAATCCCTTATCTTGGCATTT 66
DB 49474 AGATTACTTAAACAAGAAACCACTGGAAGAAAGAAATGAAATCCCTTATCTTGGCATTT 49533
QY 67 TTGGGTGGTGTCACTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGATGGCATC 126
DB 49534 TTGGGTGGTGTCACTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGATGGCATC 49593
QY 127 TCTAAGAGACCTTTTGAAGAAATTAAGAAATATAGCAAGCTGTGAGATGTTGCTAAA 186
DB 49594 TCTAAGAGACCTTTTGAAGAAATTAAGAAATATAGCAAGCTGTGAGATGTTGCTAAA 49653
QY 187 GCATATCAACCACTGCTTTATGTTAAGCCCAAGACAGATCTTATGAGCCATTGGCA 246
DB 49654 GCATATCAACCACTGCTTTATGTTAAGCCCAAGACAGATCTTATGAGCCATTGGCA 49713
QY 247 CTCTGGTGTATCTGTTGACCACTGAGTGGCTCCAGAACTTAGAAAAAGCCATC 306
DB 49714 CTCTGGTGTATCTGTTGACCACTGAGTGGCTCCAGAACTTAGAAAAAGCCATC 49773
QY 307 CAATATTATGACCAAAAAGCTGACCAAGATGGCTGGAAGAAAGTTCACTGGAGCCAGTG 366
DB 49774 CAATATTATGACCAAAAAGCTGACCAAGATGGCTGGAAGAAAGTTCACTGGAGCCAGTG 49833
QY 367 AGAATACCCCACTGGGAGAGGGGAGAAAGATCAGCTGTATGCTGGAGCCCAAGATTGAT 426
DB 49834 AGAATACCCCACTGGGAGAGGGGAGAAAGATCAGCTGTATGCTGGAGCCCAAGATTGAT 49893
QY 427 AAGATACCACTGCTGGCTTTGGCAGACAGATTGGGACTCCTCCAGAAAGGCAAT 480
DB 49894 AAGATACCACTGCTGGCTTTGGCAGACAGATTGGGACTCCTCCAGAAAGGCAAT 49947

```

RESULT 14  
 AC010859 171301 bp DNA linear PRI 24-AUG-2001  
 LOCUS Homo sapiens chromosome , clone RP11-3D19, complete sequence.  
 DEFINITION AC010859  
 ACCESSION AC010859.5 GI:9581940  
 VERSION HTG.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 171301)  
 AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Homo sapiens chromosome, clone RP11-3D19  
 Unpublished  
 2 (bases 1 to 171301)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
 Baldwin,J., Barina,N., Beckerly,R., Boguslavsky,L., Bouckgalter,B.,  
 Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,  
 Cooke,P., Dearrellano,K., Dewar,K., Domingo,M., Donegan,L., Doyle,M.,  
 Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,  
 Galagan,J., Gardyna,S., Grant,G., Hages,B., Hearford,A., Horton,L.,  
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
 Lehoczy,J., Lieu,C., Locke,K., MacDonald,P., Marquis,N.,  
 McKernan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,  
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,J.,  
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
 Wyman,D., Ye,W.-J., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (24-SEP-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 171301)  
 REFERENCE  
 AUTHORS  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barina,N., Bastien,V., Boguslavsky,L., Bouckgalter,B.,  
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,  
 Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
 Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hages,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
 Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K.,  
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,  
 Maclean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C.,  
 McCarthy,M., McKernan,P., McKernan,K., McPheters,R., Meidrum,J.,  
 Menes,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
 Raymond,C., Retica,R., Riback,W., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Roselli,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,  
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodora,J.,  
 Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.-J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (24-AUG-2001) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jul 29, 2000 this sequence version replaced gi:9280767.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIRB  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L2714  
 Center clone name: 3\_D\_19  
 -----  
 FEATURES  
 source  
 1..171301  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="RP11-3D19"  
 /clone\_1lb="RPCT-11 Human Male BAC"  
 614..2685  
 /rpt\_family="L1M42"  
 2678..3238  
 /rpt\_family="L1M42"  
 3239..4754  
 /rpt\_family="L1PA3"  
 4755..7942  
 repeat\_region

repeat_region	/rpt_family="L1MA2" 9423. .9667
repeat_region	/rpt_family="MIR" complement(10826. /rpt_family="MIR" 11047. .11079
repeat_region	/rpt_family="CT-rich" complement(11080. .11525) /rpt_family="MER82" 11690. .11737
repeat_region	/rpt_family="L2" complement(12330. .12431) /rpt_family="L2" 12432. .12891
repeat_region	/rpt_family="L2" complement(12809. /rpt_family="L2" complement(13117. .13420) /rpt_family="L2" 13513. .13667
repeat_region	/rpt_family="L2" 15490. .15681
repeat_region	/rpt_family="MER3" complement(15693. /rpt_family="MER30" complement(15895. .15965) /rpt_family="MLT1D" 15968. .15996
repeat_region	/rpt_family="TGAAn" complement(16007. .16251) /rpt_family="MLT1D" complement(16256. .16677) /rpt_family="MLT1D" 16742. .16766
repeat_region	/rpt_family="TGA n" complement(16827. .17014) /rpt_family="L3" 17090. .17252
repeat_region	/rpt_family="L2" complement(17253. .17365) /rpt_family="L1ME" complement(17417. .17934) /rpt_family="L1PB2" complement(17986. .19723) /rpt_family="L1PB8" complement(19724. .20097) /rpt_family="THERIC" complement(20098. .21674) /rpt_family="THERIC-1nt" complement(21675. .22036) /rpt_family="THERIC" complement(22037. .24561) /rpt_family="L1PBa" 24607. .24639
repeat_region	/rpt_family="AT-rich" complement(24640. .24929) /rpt_family="L1PB1" 24930. .25209
repeat_region	/rpt_family="L1PB1" 25210. .25241
repeat_region	/rpt_family="CA n" complement(25242. .25322) /rpt_family="L1PB1" 25323. .25391
repeat_region	/rpt_family="TAn n" complement(25392. .25698) /rpt_family="L1PB1" complement(25708. .25923) /rpt_family="L1PB1" complement(25953. .26528) /rpt_family="L1PBa" complement(26536. .27946) /rpt_family="TTR5"

Query Match	Best Local Similarity	83.3%; Score 472.4; DB 9; Length 171301;
Matches	473; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
repeat_region	complement (27947. .28422)	
repeat_region	/rpt_family="L1PBa"	
repeat_region	complement (28456. .28739)	
repeat_region	/rpt_family="L1MB"	
repeat_region	complement (28937. .29184)	
repeat_region	/rpt_family="MER106A"	
repeat_region	complement (29467. .29698)	
repeat_region	/rpt_family="L1MB"	
repeat_region	29792. .30497	
repeat_region	/rpt_family="L2"	
repeat_region	complement (30511. .30601)	
repeat_region	/rpt_family="L3"	
repeat_region	30636. .30857	
repeat_region	/rpt_family="MERS8A"	
repeat_region	complement (30940. .31265)	
repeat_region	/rpt_family="MLT1J2"	
repeat_region	complement (31971. .32025)	
repeat_region	/rpt_family="L3"	
repeat_region	32237. .32267	
repeat_region	/rpt_family="AT_rich"	
repeat_region	complement (32529. .32626)	
repeat_region	/rpt_family="L1MB1"	
repeat_region	32630. .32724	
repeat_region	/rpt_family="MSTC"	
repeat_region	32795. .32830	
repeat_region	/rpt_family="CA)n"	
repeat_region	32832. .34531	
repeat_region	/rpt_family="L1PAL6"	
repeat_region	34547. .34645	
repeat_region	/rpt_family="MSTC"	
repeat_region	complement (34646. .35062)	
repeat_region	/rpt_family="MSTA"	
repeat_region	35063. .35303	
repeat_region	/rpt_family="MSTC"	
repeat_region	complement (35306. .35514)	
repeat_region	/rpt_family="L1MB1"	
repeat_region	complement (36126. .36297)	
repeat_region	/rpt_family="Charlie8"	
repeat_region	37082. .37141	
repeat_region	/rpt_family="RNA-1le-ATT"	
repeat_region	complement (37825. .37936)	
repeat_region	/rpt_family="L1MC4"	
repeat_region	complement (37963. .38132)	
7	AGATTATCTTAACAGAAAAACAACCTGAGAAAAAATAAATTCCTTATCTTCCGATTT	66
93166	AGATTATCTTAACAGAAAAACAACCTGAGAAAAAATAAATTCCTTATCTTCCGATTT	932255
67	TTGGTGGTGTACCTTTATCCCTGTGCTCTGGGAAGCTATATGCAAGAATGGCATTC	126
93226	TTGGTGGTGTGTACCTTTATCCCTGTGCTCTGGGAAGCTATATGCAAGAATGGCATTC	932855
127	TCTAAGAGACCTTTGAGAAATAAAGAAATAAGCAGCTGTGAGATGTTCTAAA	186
93286	TCTAAGAGACCTTTGAGAAATAAAGAAATAAGCAGCTGTGAGATGTTCTAAA	933455
187	GCAATCATCAACCTAGCTGTTTATGTGTAAGCCCGACAGACAGATCCTATGAGCGATTGGCA	246
93346	GCAATCATCAACCTAGCTGTTTATGTGTAAGCCCGACAGACAGATCCTATGAGCGATTGGCA	934050
247	CTTCTGGTGTATCTGTGTGAGCCCAACCTGAGTGGCTCCAGAAACCTAAGAAAACCATTC	306
93406	CTTCTGGTGTATCTGTGTGAGCCCAACCTGAGTGGCTCCAGAAACCTAAGAAAACCATTC	934655
307	CAAAATTAATGACAAAACCTGACGAAATGGCTGGAAGAAGTTCAACCTGAGACGACGTG	366
93466	CAAAATTAATGACAAAACCTGACGAAATGGCTGGAAGAAGTTCAACCTGAGACGACGTG	935255
367	AGAAATCCCCACTGGAGAGGGGAGAAAGATCAAGCTGTGATCTGGAGCCAAAGATTCAAT	426

Db 93526 AGAATACCCCATGGGAGAGGGAGAAATCAGCTGTGATGCTGGAGCCCAAGATTCAT 93585  
QY 427 AAGATGCCATTCCTGGGCTTTGGCAGACGATTTGGGATCTCTCCAGAGGAT 480  
Db 93586 AAGATGCCATTCCTGGGCTTTGGCAGACGATTTGGGATCTCTCCAGAGGAT 93639

RESULT 15  
AF131077  
LOCUS AF131077 1726 bp mRNA linear ROD 03-JAN-2003  
DEFINITION Rattus norvegicus liver annexin-like protein (LAL) mRNA, complete cds.  
ACCESSION AF131077 GI:7108712  
VERSION AF131077.1  
KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 1726)  
AUTHORS Della Fazio, M.A., Piobbico, D., Bartoli, D., Castelli, M., Brancorsini, S., Viola Magni, M. and Servillo, G.  
TITLE lal-1: a differentially expressed novel gene during proliferation in liver regeneration and in hepatoma cells  
JOURNAL Genes Cells 7 (11), 1183-1190 (2002)  
MEDLINE 22278398  
PubMed 12390252

REFERENCE 2 (bases 1 to 1726)  
AUTHORS Servillo, G., Della Fazio, M.A., Piobbico, D., Bartoli, D., Castelli, M., Brancorsini, S. and Viola Magni, M.  
TITLE Direct Submission  
JOURNAL Submitted (25-FEB-1999) Institute of General Pathology, University of Perugia, Policlinico Monteluce, Perugia 06100, Italy  
FEATURES  
source location/Qualifiers  
1..1726  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
1..1726  
/gene="LAL"  
63..1481  
/gene="LAL"  
/note="Involved in liver regeneration; contains unique annexin domain at the N-terminus"  
/codon\_start=1  
/product="liver annexin-like protein"  
/protein\_id="AAF36518.1"  
/db\_xref="GI:7108713"  
translation="MRPLFLFVAVVHLFSLGSKATYKSGVSORTFOEIKERIANYE  
DVAKATINLAVTKYONRSTYERGLVDIVGPRLSGSNNLEKAIQIMTQNDGLN  
VHLEQVRI PHMERGESAVVWVPRIRHKLALIGLGSIGTPEGITAEVLVASFVLEQ  
RRASEARGKIVVNPYTDYGTQYERBGAFAKAGVAVSLRVAESFISVPHGT  
HOGYODGVPRIPACITIEDAEMMSRTASRGDKIVILKMGATKYPDTSFNTVAEIT  
GSXKPEEVIVVSGHLSMDVGGALDDGGAFLSMELSLVKDLGLPKRTLRVLMET  
AEBGCGVAGASYIELHKANISKISLVWEADSGTFLPGLQPTGSDKRAAIKREYMSLI  
QPLNITVFNDABGTDINFWIQAGVPGASLKDILYKTFPHSHSGDTMTADPKOMNV  
AAAVWAVVAVVADMEELPRS"

BASE COUNT 474 a 385 c 428 g 439 t  
ORIGIN

Query Match 70.5%; Score 400; DB 10; Length 1726;  
Best Local Similarity 82.8%; Pred. No. 2.4e-100;  
Matches 457; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 15 TTAACAAGAAACCACTGAGAAAAAATAAATTCCTATCTTCGCAATTTTCGGTGG 74  
Db 35 TGACCAAGAAAGAAAGAACTAGGACATAGGTTCTTTCTTCCTGTTGCTGT 94  
QY 75 TGTTCACCTTTTATTCCTGCTGTGGGAAAGCTATATGCAAGATGGCATCTTAAG 134  
Db 95 TGTTCACCTTTTCTTCCTGCTGTGGGAAAGCTATATGCAAGATGGTGTCTCAGCG 154

QY 135 GACTTTGAAGAAATATAAGAAATAGCCACCTGTGAGATGTTGCTAAAGCAATCAT 194  
Db 155 ACATTTCAAGAAATATAAGAAATAGCCACCTGTGAGATGTTGCTAAAGCAATCAT 214  
QY 195 CAACCTAGCTGTTTATGATTAAGCCAGAAACAGATCTATAGCGATTTGCACTTCTG 254  
Db 215 CAACCTAGCTGTTTATGATTAAGAAATATCCAGAACCGGTGATGATGCTTTGGGACTTCTAGT 274  
QY 255 TGATACCTGTGAGAACCCAGACTGATGCTGCTCCAGAAACCTGAAAGAACCTCCAAATAT 314  
Db 275 TGATACCTGTGAGAACCCAGACTGATGCTGCTCCAGAAACCTGAAAGAACCTCCAAATAT 334  
QY 315 GTACCAAAACCTGAGCAAGATGAGCTGTGAGAAAGTTCACTGAGCCAGTGAATATACC 374  
Db 335 GTACCAAAACCTGAGCAAGATGAGCTGTGAGAAAGTTCACTGAGCCAGTGAATATACC 394  
QY 375 CCACTGGAGAGGAGGAGAGAAATCACTGTGATGCTGTGAGCCAGAAATTCATAGATAGC 434  
Db 395 TCACTGGAGAGGAGGAGGAGAAATCTGCAGTGATGCTGTGAGCCAGAAATTCATAGATAGC 454  
QY 435 CATTCTGGCTCTTGGCAGCAGCATTTGGGACTCTCTCCAGAAAGCATTCACAGAAATGCT 494  
Db 455 TATTTTAGGCTTGGCGGAGCATTTGGGACTCTCTCCAGAAAGCATTCACAGAAATGCT 514  
QY 495 GGTGTGACCTCTTTCATGATGACCTGAGAGAGGCTCAGAAAGAGAGGAAATTTGT 554  
Db 515 GGTGTGACCTCTTTCATGATGACCTGAGAGAGGCTCAGAAAGAGAGGAAATTTGT 574  
QY 555 TGTATTATACCA 566  
Db 575 TGTATTATACCA 586

Search completed: December 22, 2003, 20:22:27  
Job time : 2052.56 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:18:08 ; Search time 6681.9 Seconds  
(without alignments)  
11332.658 Million cell updates/sec

Title: US-09-745-763-35

Perfect score: 1851

Sequence: 1 GGCTAGCGCCGAGCTTAGT.....CTGAAAAAAAAAAAAAAAAA 1851

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

```

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sv:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1851	100.0	1851	6 BD106411	BD106411 Secreted
2	1839.8	99.4	1928	6 BC020689	BC020689 Homo sapi
3	1820.4	98.3	1860	6 BD127520	BD127520 Primer fo
4	1820.4	98.3	1860	9 AK075132	AK075132 Homo sapi
5	1783.2	96.3	1860	6 AX006440	AX006440 Sequence
6	1758.6	95.0	1794	9 AF119386	AF119386 Homo sapi
7	1751	94.6	1778	6 BD105816	BD105816 Secretory
8	1465.6	79.2	1472	6 AF107834	AF107834 Homo sapi
9	1164.8	62.9	1796	10 BC037067	BC037067 Mus muscu
10	1145	61.9	1778	6 AX400829	AX400829 Sequence
11	1145	61.9	1778	10 AF097723	AF097723 Rattus no
12	1143.8	61.8	1726	10 AF131077	AF131077 Rattus no
13	1118.4	60.4	1596	10 AF107835	AF107835 Mus muscu
14	1077.8	58.2	1716	10 AF009513	AF009513 Mus muscu
15	569	30.7	609	6 BD059610	BD059610 Secreted
16	568.2	30.7	642	6 BD125210	BD125210 Primer fo
17	568.2	30.7	642	6 BD126476	BD126476 Primer fo
18	484.4	26.2	152176	9 AP003117	AP003117 Homo sapi
19	484.4	26.2	166050	2 AC013817	AC013817 Homo sapi
20	484.4	26.2	177008	2 AC011175	AC011175 Homo sapi
21	472.4	25.5	156702	2 AC084223	AC084223 Homo sapi
22	472.4	25.5	157927	9 AP006278	AP006278 Homo sapi
23	472.4	25.5	171301	9 AC010859	AC010859 Homo sapi
24	421.4	22.8	423	9 AF107833	AF107833 Homo sapi
25	410.6	22.2	492	6 BD125961	BD125961 Primer fo
26	331	17.9	417	10 AF107832	AF107832 Mus muscu
27	327.4	17.7	167883	2 AC121026	AC121026 Rattus no
28	327.4	17.7	218874	2 AC110103	AC110103 Rattus no
29	324.4	17.5	360	6 BD105812	BD105812 Secretory
30	324.4	17.5	360	6 E62850	E62850 Secretory p
31	323.6	17.5	214287	2 AC129951	AC129951 Mus muscu
32	310.2	16.8	317	6 BD076934	BD076934 5' EST of
33	280.8	15.7	314	6 BD058492	BD058492 Secreted
34	284.6	15.4	462	6 BD105811	BD105811 Secretory
35	284.6	15.4	462	6 E62849	E62849 Secretory p
36	233	12.6	234	6 AX411237	AX411237 Sequence
37	233	12.6	234	11 G07404	G07404 human SRS W
38	212.2	11.5	144703	9 AP003112	AP003112 Homo sapi
39	205	11.1	171035	9 AP003111	AP003111 Homo sapi
40	174.4	9.4	214287	2 AC129951	AC129951 Mus muscu
41	174.2	9.4	240178	2 AC094173	AC094173 Rattus no
42	172.8	9.3	203469	2 AC123654	AC123654 Mus muscu
43	171.2	9.2	218874	2 AC110103	AC110103 Rattus no
44	163.4	8.8	288136	2 AC120489	AC120489 Rattus no
45	154.4	8.3	238034	2 AC123681	AC123681 Mus muscu

## ALIGNMENTS

```

RESULT 1
LOCUS BD106411 1851 bp DNA linear PAT 18-SEP-2002
DEFINITION Secreted proteins and polynucleotides encoding them.
ACCESSION BD106411
VERSION BD106411.1 GI:23201229
KEYWORDS JP 2002503955-A/2.
SOURCE Chlamydia sp.
ORGANISM Chlamydia sp.
REFERENCE 1 (bases 1 to 1851)
Bacteria: Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
Jacobae, K., McCoy, J.M., Lavallie, E.R., Racie, L.A., Merberg, D.,
Treacy, M., Spaulding, V., and Agostino, M.J.
Secreted proteins and polynucleotides encoding them
Patent: JP 2002503955-A 2 05-FEB-2002;
JOURNAL

```

GENETICS INSTITUTE INC  
 PN JP 2002503955-A/2  
 PD 05-FEB-2002  
 PF 20-MAR-1998 JP 1998545874  
 PR 21-MAR-1997 US 08/822167, 19-MAR-1998 US 09/044466 PI  
 KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI  
 DAVID MERBERG,  
 PI MAURICE TREACY, VIKKI SPAILDING, MICHAEL J AGOSTINO PC  
 CI2N15/12.C07K14/47.A61K38/17  
 CC Strandedness: Double;  
 CC Topology: Linear;  
 FH Key Location/Qualifiers.

FEATURES  
 Source 1..1851  
 /organism="Chlamydia sp."  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:35827"

BASE COUNT 531 a 413 c 438 g 469 t  
 ORIGIN

Query Match 100.0%; Score 1851; DB 6; Length 1851;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1851; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 GGCTAGGCGGAGCTTGTCTCTGGAGCCGCTCCGTCGCGCCGTCAGAGCCGCCCTTA
DB 1 GGCTAGGCGGAGCTTGTCTCTGGAGCCGCTCCGTCGCGCCGTCAGAGCCGCCCTTA 60
OY 61 TCAGATTATCTTAAACAAGAAAACAAGTGAAGAAAAAATGAAATTCCTTATCTTCGCAT
DB 61 TCAGATTATCTTAAACAAGAAAACAAGTGAAGAAAAAATGAAATTCCTTATCTTCGCAT 120
OY 121 TTTTCGGGTGGTCACTTTTATCCCTGTGCTCTGGAGAAAGCTTATGAGAAATGGCA
DB 121 TTTTCGGGTGGTCACTTTTATCCCTGTGCTCTGGAGAAAGCTTATGAGAAATGGCA 180
OY 121 TTTTCGGGTGGTCACTTTTATCCCTGTGCTCTGGAGAAAGCTTATGAGAAATGGCA 180
DB 121 TTTTCGGGTGGTCACTTTTATCCCTGTGCTCTGGAGAAAGCTTATGAGAAATGGCA 180
OY 181 TCTCTAAGAGAGCTTTTGAAGAAATGAAGAAATAGCAGCTGTGAGATGTGGTA
DB 181 TCTCTAAGAGAGCTTTTGAAGAAATGAAGAAATAGCAGCTGTGAGATGTGGTA 240
OY 181 TCTCTAAGAGAGCTTTTGAAGAAATGAAGAAATAGCAGCTGTGAGATGTGGTA 240
DB 181 TCTCTAAGAGAGCTTTTGAAGAAATGAAGAAATAGCAGCTGTGAGATGTGGTA 240
OY 241 AAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCCAAGACAGATCTTATGAGCATTTGG
DB 241 AAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCCAAGACAGATCTTATGAGCATTTGG 300
OY 241 AAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCCAAGACAGATCTTATGAGCATTTGG 300
DB 241 AAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCCAAGACAGATCTTATGAGCATTTGG 300
OY 301 CACTTCGTGGTGAATCTGTGGACCCAGACAGATGGCTCCAAAGAACTTAAAGAAAGCA
DB 301 CACTTCGTGGTGAATCTGTGGACCCAGACAGATGGCTCCAAAGAACTTAAAGAAAGCA 360
OY 301 CACTTCGTGGTGAATCTGTGGACCCAGACAGATGGCTCCAAAGAACTTAAAGAAAGCA 360
DB 301 CACTTCGTGGTGAATCTGTGGACCCAGACAGATGGCTCCAAAGAACTTAAAGAAAGCA 360
OY 361 TCCAAATATATGACCAAAACCTGACAGCAAGATGGCTGGAAGAAATTCACCTGAGCCAG
DB 361 TCCAAATATATGACCAAAACCTGACAGCAAGATGGCTGGAAGAAATTCACCTGAGCCAG 420
OY 361 TCCAAATATATGACCAAAACCTGACAGCAAGATGGCTGGAAGAAATTCACCTGAGCCAG 420
DB 361 TCCAAATATATGACCAAAACCTGACAGCAAGATGGCTGGAAGAAATTCACCTGAGCCAG 420
OY 421 TGAGATATACCCCACTGGAGAGAGGAGAAATCAAGCTGTGATGCTGAGCCCAAGATTC
DB 421 TGAGATATACCCCACTGGAGAGAGGAGAAATCAAGCTGTGATGCTGAGCCCAAGATTC 480
OY 421 TGAGATATACCCCACTGGAGAGAGGAGAAATCAAGCTGTGATGCTGAGCCCAAGATTC 480
DB 421 TGAGATATACCCCACTGGAGAGAGGAGAAATCAAGCTGTGATGCTGAGCCCAAGATTC 480
OY 481 ATAAGATGCGCATCTGGGCTCTTGGCAGCAGATTTGGACCTCTCAAGAGCAATTAAG
DB 481 ATAAGATGCGCATCTGGGCTCTTGGCAGCAGATTTGGACCTCTCAAGAGCAATTAAG 540
OY 481 ATAAGATGCGCATCTGGGCTCTTGGCAGCAGATTTGGACCTCTCAAGAGCAATTAAG 540
DB 481 ATAAGATGCGCATCTGGGCTCTTGGCAGCAGATTTGGACCTCTCAAGAGCAATTAAG 540
OY 541 CAGAAGTCTGAGTGAACCTCTTTCAGATGAACCTGACAGAGAGGAGCTCAGAGCAAG
DB 541 CAGAAGTCTGAGTGAACCTCTTTCAGATGAACCTGACAGAGAGGAGCTCAGAGCAAG 600
OY 541 CAGAAGTCTGAGTGAACCTCTTTCAGATGAACCTGACAGAGAGGAGCTCAGAGCAAG 600
DB 541 CAGAAGTCTGAGTGAACCTCTTTCAGATGAACCTGACAGAGAGGAGCTCAGAGCAAG 600
OY 601 GGAAGATTGTTTATTAACAACCTTATCATCACTCAAGAGAGCGGTGCAATACCGAA
DB 601 GGAAGATTGTTTATTAACAACCTTATCATCACTCAAGAGAGCGGTGCAATACCGAA 660
OY 601 GGAAGATTGTTTATTAACAACCTTATCATCACTCAAGAGAGCGGTGCAATACCGAA 660
DB 601 GGAAGATTGTTTATTAACAACCTTATCATCACTCAAGAGAGCGGTGCAATACCGAA 660
OY 661 CGCAGGAGGCGGTGGAAGCTGCCAAGGTGGGCTTTGGCATCTTCATTCGATCCGTGG
DB 661 CGCAGGAGGCGGTGGAAGCTGCCAAGGTGGGCTTTGGCATCTTCATTCGATCCGTGG 720
OY 661 CGCAGGAGGCGGTGGAAGCTGCCAAGGTGGGCTTTGGCATCTTCATTCGATCCGTGG 720
DB 661 CGCAGGAGGCGGTGGAAGCTGCCAAGGTGGGCTTTGGCATCTTCATTCGATCCGTGG 720
OY 721 CCTCCTTCTCATCTAAGTCTCAACAGGTATTCAGGAATACAGAGATGGCGTCCCA
DB 721 CCTCCTTCTCATCTAAGTCTCAACAGGTATTCAGGAATACAGAGATGGCGTCCCA 780

```

```

DB 721 CCTCCTTCTCATCTAAGTCTCAACAGGTATTCAGGAATACAGAGATGGCGTCCCA 780
OY 781 AAATTCACAAGCCTGTATTAACGTGGAAGATGAGAAATGATGTCAAGATGGCTTCTC
DB 781 AAATTCACAAGCCTGTATTAACGTGGAAGATGAGAAATGATGTCAAGATGGCTTCTC 840
OY 841 ATGGGATCAAAATGTGATTCAGTAAAGATGGGGGCAAGAACTTACCAGATCTGATTT
DB 841 ATGGGATCAAAATGTGATTCAGTAAAGATGGGGGCAAGAACTTACCAGATCTGATTT 900
OY 841 ATGGGATCAAAATGTGATTCAGTAAAGATGGGGGCAAGAACTTACCAGATCTGATTT 900
DB 841 ATGGGATCAAAATGTGATTCAGTAAAGATGGGGGCAAGAACTTACCAGATCTGATTT 900
OY 901 CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTTGTACTGTCA
DB 901 CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTTGTACTGTCA 960
OY 901 CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTTGTACTGTCA 960
DB 901 CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATATCCAGAACAGGTTGTACTGTCA 960
OY 961 GTGACATCTGGAACAGCTGGAGATTTGGGAGGCTGCATGATGATGGCGGTGGACCT
DB 961 GTGACATCTGGAACAGCTGGAGATTTGGGAGGCTGCATGATGATGGCGGTGGACCT 1020
OY 961 GTGACATCTGGAACAGCTGGAGATTTGGGAGGCTGCATGATGATGGCGGTGGACCT 1020
DB 961 GTGACATCTGGAACAGCTGGAGATTTGGGAGGCTGCATGATGATGGCGGTGGACCT 1020
OY 1021 TTATATCATGGAAGACACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTC
DB 1021 TTATATCATGGAAGACACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTC 1080
OY 1021 TTATATCATGGAAGACACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTC 1080
DB 1021 TTATATCATGGAAGACACTCTCACTTATTAAGATCTTGGGCTGCTCCAAAGAGACTC 1080
OY 1081 TGCGGCTGTGCTCTGAGACTGACAGAGAACAGGTGAGATTTGCTTCCAGTATTATC
DB 1081 TGCGGCTGTGCTCTGAGACTGACAGAGAACAGGTGAGATTTGCTTCCAGTATTATC 1140
OY 1081 TGCGGCTGTGCTCTGAGACTGACAGAGAACAGGTGAGATTTGCTTCCAGTATTATC 1140
DB 1081 TGCGGCTGTGCTCTGAGACTGACAGAGAACAGGTGAGATTTGCTTCCAGTATTATC 1140
OY 1141 AGTTACACAAGTAAATATTTCCAACTACAGCTCTGTGATGAGATCTGACCCAGAACT
DB 1141 AGTTACACAAGTAAATATTTCCAACTACAGCTCTGTGATGAGATCTGACCCAGAACT 1200
OY 1141 AGTTACACAAGTAAATATTTCCAACTACAGCTCTGTGATGAGATCTGACCCAGAACT 1200
DB 1141 AGTTACACAAGTAAATATTTCCAACTACAGCTCTGTGATGAGATCTGACCCAGAACT 1200
OY 1201 TCTTACCACATGGGCTCAATATCACTGAGAGTGAAGAGCCAGGGCCATCATGAGAGAG
DB 1201 TCTTACCACATGGGCTCAATATCACTGAGAGTGAAGAGCCAGGGCCATCATGAGAGAG 1260
OY 1201 TCTTACCACATGGGCTCAATATCACTGAGAGTGAAGAGCCAGGGCCATCATGAGAGAG 1260
DB 1201 TCTTACCACATGGGCTCAATATCACTGAGAGTGAAGAGCCAGGGCCATCATGAGAGAG 1260
OY 1261 TTTAGACCTGTGCTGACAGCCCTCAATATCACTGAGTCTGAGCCATGAGAGAGAGAG
DB 1261 TTTAGACCTGTGCTGACAGCCCTCAATATCACTGAGTCTGAGCCATGAGAGAGAGAG 1320
OY 1261 TTTAGACCTGTGCTGACAGCCCTCAATATCACTGAGTCTGAGCCATGAGAGAGAGAG 1320
DB 1261 TTTAGACCTGTGCTGACAGCCCTCAATATCACTGAGTCTGAGCCATGAGAGAGAGAG 1320
OY 1321 ACATCAACTTTTGAATCAAGCTGAGAGTGGCTGAGAGCAGCTTATGATGATTAACA
DB 1321 ACATCAACTTTTGAATCAAGCTGAGAGTGGCTGAGAGCAGCTTATGATGATTAACA 1380
OY 1321 ACATCAACTTTTGAATCAAGCTGAGAGTGGCTGAGAGCAGCTTATGATGATTAACA 1380
DB 1321 ACATCAACTTTTGAATCAAGCTGAGAGTGGCTGAGAGCAGCTTATGATGATTAACA 1380
OY 1381 AGTATTTCTTCTTCACTACCTCCAGAGAGACACAGATGATGATGATGATGATGATGATG
DB 1381 AGTATTTCTTCTTCACTACCTCCAGAGAGACACAGATGATGATGATGATGATGATGATG 1440
OY 1381 AGTATTTCTTCTTCACTACCTCCAGAGAGACACAGATGATGATGATGATGATGATGATG 1440
DB 1381 AGTATTTCTTCTTCACTACCTCCAGAGAGACACAGATGATGATGATGATGATGATGATG 1440
OY 1441 TGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
DB 1441 TGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
OY 1441 TGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
DB 1441 TGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
OY 1501 TGCTGCTAGGTCCTTAAGAAACAGTAAGAAAGAAAGCTTCAATGCTTCTGGCCAGGAATC
DB 1501 TGCTGCTAGGTCCTTAAGAAACAGTAAGAAAGAAAGCTTCAATGCTTCTGGCCAGGAATC 1560
OY 1501 TGCTGCTAGGTCCTTAAGAAACAGTAAGAAAGAAAGCTTCAATGCTTCTGGCCAGGAATC 1560
DB 1501 TGCTGCTAGGTCCTTAAGAAACAGTAAGAAAGAAAGCTTCAATGCTTCTGGCCAGGAATC 1560
OY 1561 CTGGGCTGCAACTTTGGAAGAACTCTCTTCACTAATCAATTTGATCCAAATTCATCTTCA
DB 1561 CTGGGCTGCAACTTTGGAAGAACTCTCTTCACTAATCAATTTGATCCAAATTCATCTTCA 1620
OY 1561 CTGGGCTGCAACTTTGGAAGAACTCTCTTCACTAATCAATTTGATCCAAATTCATCTTCA 1620
DB 1561 CTGGGCTGCAACTTTGGAAGAACTCTCTTCACTAATCAATTTGATCCAAATTCATCTTCA 1620
OY 1621 AAGCAAACTCTATTTCAATGCTTCTGTTATTAATCTTGTGATGATCTTCCAAATTCCT
DB 1621 AAGCAAACTCTATTTCAATGCTTCTGTTATTAATCTTGTGATGATCTTCCAAATTCCT 1680
OY 1621 AAGCAAACTCTATTTCAATGCTTCTGTTATTAATCTTGTGATGATCTTCCAAATTCCT 1680
DB 1621 AAGCAAACTCTATTTCAATGCTTCTGTTATTAATCTTGTGATGATCTTCCAAATTCCT 1680
OY 1681 GATTCTGAAAAAAGAAATCAATTCCTCCCTCCCTCCCAACATGAAATCAACATATGTTA
DB 1681 GATTCTGAAAAAAGAAATCAATTCCTCCCTCCCTCCCAACATGAAATCAACATATGTTA 1740
OY 1681 GATTCTGAAAAAAGAAATCAATTCCTCCCTCCCTCCCAACATGAAATCAACATATGTTA 1740
DB 1681 GATTCTGAAAAAAGAAATCAATTCCTCCCTCCCTCCCAACATGAAATCAACATATGTTA 1740
OY 1741 GGGATTCAGTGGGGGATTTCTTATATCACTCTTAAATAAATGTTTCCATTTTAA
DB 1741 GGGATTCAGTGGGGGATTTCTTATATCACTCTTAAATAAATGTTTCCATTTTAA 1800
OY 1741 GGGATTCAGTGGGGGATTTCTTATATCACTCTTAAATAAATGTTTCCATTTTAA 1800
DB 1741 GGGATTCAGTGGGGGATTTCTTATATCACTCTTAAATAAATGTTTCCATTTTAA 1800
OY 1801 AGTAAACACTTAATTAATTTTGGAGATCTCTGAAAAAATGAAAAAATGAAAAAATGAAAAA
DB 1801 AGTAAACACTTAATTAATTTTGGAGATCTCTGAAAAAATGAAAAAATGAAAAAATGAAAAA 1851

```



Db 1801 AGTAAACATTAAATTTTGGAGAGATCTGTGAAAAA 1851

RESULT 2  
BC020689 1928 bp mRNA linear PRI 22-JAN-2002  
DEFINITION Homo sapiens, plasma glutamate carboxypeptidase, clone MGC:22418  
IMAGE:4251802, mRNA, complete cds.  
ACCESSION BC020689  
VERSION BC020689.1 GI:18088383  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
Strausberg, R.  
REFERENCE 1 (bases 1 to 1928)  
TITLE Direct Submission  
AUTHORS Submitted (03-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
JOURNAL NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
REMARK Contact: MGC help desk  
COMMENT Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC/DCTD/DRP  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcdpaxil.stanford.edu](mailto:mcdpaxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>  
Series: IRAL Plate: 35 Row: F Column: 13  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7706386.  
Location/Qualifiers  
1. 1928  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="LocusID:10404"  
/db\_xref="taxon:9606"  
/clone="MGC:22418 IMAGE:4251802"  
/issue\_type="Skin, melanocytic melanoma, high MDR."  
/clone\_id="NIH\_MGC\_62"  
/lab\_host="DH10B"  
/note="vector: pDNR-LIB"  
167.1585  
/codon\_start=1  
/product="plasma glutamate carboxypeptidase"  
/protein\_id="AAH20689.1"  
/db\_xref="GI:18088384"  
/translation="MKRLIRAFKFGVHLISCSGKAIQKNGISRTFEIKERIASGCVAKATINLAVYGAQNRSTYERLAVDTGPRISGKNLEKAIQIQTQNIQDGLK VHLERPAIPIHERGESAVLMEPRIHAIIGLSSIGTPEEGITAVLVVTSDELQ RASERAGKIVVNPQYINYSRTQVOTGQVNAKVALASLRSVASFSISPHNG IOEVODVPIKIPACTIYEDAEEMSRASHGIKIVIQKMAKTYPTDSFNTYABIT GSXKPEOVVLVSHLSDMDVGGAMDGGGAFISWEALSILKIDGLPKRTLRVMT AERGGVGAQRYVOLHAKVINSYLSNESDAGTFLPGLGLOPTSEKRAIMBERMSLI QPLATITVLSHSGSTIDNFMIOAGVPQASLDDLYKTFPHSHSGDTTVMDFKQNNV AAAVAVVSYVADMEMLPRS"

BASE COUNT 554 a 435 c 466 g 473 t  
ORIGIN

Query Match 99.4%; Score 1839.8; DB 9; Length 1928;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1844; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GGCTAGGCGCGACCTTAGTCTCTGGAGAGCCGCTTCCTCGCGCGGTACAGAGCCGCCCTA 60

Db 69 GGCTAGGCGCGAGCTTAGTCTCTGGAGAGCCGCTTCCTCGCGCGGTACAGAGCCGCCCTA 128

QY 61 TCGATTATCTTAAACAAAGAAACCACTGGAAGAAAAAATTCCTATCTGGCAT 120

Db 129 TCGATTATCTTAAACAAAGAAACCACTGGAAGAAAAAATTCCTATCTGGCAT 188

QY 121 TTTTCGGTGTGTTCACCTTTTATCCCTGTGCTCTGGAAAGCTATATGCAAGATGCA 180

Db 189 TTTTCGGTGTGTTCACCTTTTATCCCTGTGCTCTGGAAAGCTATATGCAAGATGCA 248

QY 181 TCTTAAAGAGACTTTTGAAGAAATTAAGAAATATGACCTGTGAGATGTGCTA 240

Db 249 TCTTAAAGAGACTTTTGAAGAAATTAAGAAATATGACCTGTGAGATGTGCTA 308

QY 241 AAGCAATCAACCACTTGTGTTATGCTTAAGCCCAAGAACATCTATGACGATTTGG 300

Db 309 AAGCAATCAACCACTTGTGTTATGCTTAAGCCCAAGAACATCTATGACGATTTGG 368

QY 301 CACTTGTGTGATCTGTTGAGACCCAGACTGAGTGGCTCAAGAACTTGAAGAAAGCCA 360

Db 369 CACTTGTGTGATCTGTTGAGACCCAGACTGAGTGGCTCAAGAACTTGAAGAAAGCCA 428

QY 361 TCCAAATTATGTACCAAAACCTGACAGACAGATGAGTGGAGAAAGTTCACTGAGGCGAG 420

Db 429 TCCAAATTATGTACCAAAACCTGACAGACAGATGAGTGGAGAAAGTTCACTGAGGCGAG 488

QY 421 TGAGAAATACCCCACTGGGAGAGAGGAGAGAAATCACTGTGATGCTGAGACCAAGATTC 480

Db 489 TGAGAAATACCCCACTGGGAGAGAGGAGAGAAATCACTGTGATGCTGAGACCAAGATTC 548

QY 481 ATAGATAGCCATCCCTGGGATCTTGGAGAGCAATGGGACCTCCCAAGAGGCAATTACAG 540

Db 549 ATAGATAGCCATCCCTGGGATCTTGGAGAGCAATGGGACCTCCCAAGAGGCAATTACAG 608

QY 541 CAGAAATTCTGTGTGATGACCTCTTTCGATGACCTGACAGAGAGGCTCAGAAAGCAAGAG 600

Db 609 CAGAAATTCTGTGTGATGACCTCTTTCGATGACCTGACAGAGAGGCTCAGAAAGCAAGAG 668

QY 601 GGAAGATTTGTTTATTAACCAACTTATCACTCACTCAAGAGCGGTGCATACCGAA 660

Db 669 GGAAGATTTGTTTATTAACCAACTTATCACTCACTCAAGAGCGGTGCATACCGAA 728

QY 661 CGCAGGGGGCGGTGGAAGCTGCCAAGGTGGGGGCTTGGCATCTCATTCGATCCGTTGG 720

Db 729 CGCAGGGGGCGGTGGAAGCTGCCAAGGTGGGGGCTTGGCATCTCATTCGATCCGTTGG 788

QY 721 CTTCTTCTCATCTACAGTCTCTCACACAGATATTCAAGAAATCAAGATGGCGTGCCTCA 780

Db 789 CTTCTTCTCATCTACAGTCTCTCACACAGATATTCAAGAAATCAAGATGGCGTGCCTCA 848

QY 781 AATTCCACAGCCTGTATTACGCTGAGAGATGACAGAAATGATGTCAAGAAATGGCTTTC 840

Db 849 AGATTCCACAGCCTGTATTACGCTGAGAGATGACAGAAATGATGTCAAGAAATGGCTTTC 908

QY 841 ATGGGATCAAAATTTGATTCAGTCAAGATGGGGGCAAGACCTTACCAGATCTGATTT 900

Db 909 ATGGGATCAAAATTTGATTCAGTCAAGATGGGGGCAAGACCTTACCAGATCTGATTT 968

QY 901 CTTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATCCAGAACAGTTGTACTGTCA 960

Db 969 CTTTCAACACTGTAGCAGAGATCACTGGGAGCAAAATATCCAGAACAGTTGTACTGTCA 1028

QY 961 GTGACATCTGACAGCTGGAGATGTTGGCAGAGGTGTCATGATGATGCGGTGGAGCCT 1020

Db 1029 GTGACATCTGACAGCTGGAGATGTTGGCAGAGGTGTCATGATGATGCGGTGGAGCCT 1088

QY 1021 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTGTGGGCGGTCCAAAGAGGAGCTC 1080

Db 1089 TTATATCATGGGAAGCACTCTCACTTATTAAGATCTGTGGGCGGTCCAAAGAGGAGCTC 1148

QY 1081 TGCAGCTGTGCTCTGTGACTGAGAGAAAGCAAGGTGAGTTGTGCTTCAGATATTATC 1140

Db 1149 TCGCGCTGCTGCTGCACTGCAAGAAACAAGTGAGTTGCTTCAGTATATC 1208  
 Oy 1141 AGTTACACAGGTAATATTTTCACTACAGTCTGATGAGTCTGACGAGGAACT 1200  
 Db 1209 AGTTACACAGGTAATATTTTCACTACAGTCTGATGAGTCTGACGAGGAACT 1268  
 Oy 1201 TCTTACCACTGGGCTGCAATTTCACTGCAAGTGAAGGCAAGGCAATCAATGAGAGG 1260  
 Db 1269 TCTTACCACTGGGCTGCAATTTCACTGCAAGTGAAGGCAAGGCAATCAATGAGAGG 1328  
 Oy 1261 TTTATAGCTCTGTCAGAGCCCTCAATATCACTCAGGTCTTGAGCCATGAGAGGAGAG 1320  
 Db 1329 TTTATAGCTCTGTCAGAGCCCTCAATATCACTCAGGTCTTGAGCCATGAGAGGAGAG 1388  
 Oy 1321 ACATCACTTTTGTGATCCAGAGTGAAGTCTGAGCCAGTCTATCTTATGACTATACA 1380  
 Db 1389 ACATCACTTTTGTGATCCAGAGTGAAGTCTGAGCCAGTCTATCTTATGACTATACA 1448  
 Oy 1381 AGTATTTCTTCTTCATCACTCCACGAGACACCATGATCTGTCAATGATCCAAAGCAGA 1440  
 Db 1449 AGTATTTCTTCTTCATCACTCCACGAGACACCATGATCTGTCAATGATCCAAAGCAGA 1508  
 Oy 1441 TGAATGCTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGACATGAGAGAA 1500  
 Db 1509 TGAATGCTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGACATGAGAGAA 1568  
 Oy 1501 TGCTGCTAGGCTCTAGAAACGTAAGAAAGTTTCAATGCTTCTGCGCAGGAATC 1560  
 Db 1569 TGCTGCTAGGCTCTAGAAACGTAAGAAAGTTTCAATGCTTCTGCGCAGGAATC 1628  
 Oy 1561 CTGGGCTGCAACTTTGAAAACTCCTCTTCACATACAAATTTCAATTCATCTCA 1620  
 Db 1629 CTGGGCTGCAACTTTGAAAACTCCTCTTCACATACAAATTTCAATTCATCTCA 1688  
 Oy 1621 AAGCACTATTTTCAATGCTTCTGTTATATCTTCTTGATATCTTCCAAATCTCT 1680  
 Db 1689 AAGCACTATTTTCAATGCTTCTGTTATATCTTCTTGATATCTTCCAAATCTCT 1748  
 Oy 1681 GATTTAGAAAAAGAAATCATCTCCCTCCCTCCACACATACATCAATATGAGTA 1740  
 Db 1749 GATTTAGAAAAAGAAATCATCTCCCTCCCTCCACACATACATCAATATGAGTA 1808  
 Oy 1741 GGAATTAAGTGGGGCAATTTCTTATATCACTCTTAAAAAATTTGTTTCACTTTAA 1800  
 Db 1809 GGAATTAAGTGGGGCAATTTCTTATATCACTCTTAAAAAATTTGTTTCACTTTAA 1868  
 Oy 1801 AGTAACCTTATATATTTTGGAGATCTCTGAAAAAATTTTAAAAA 1851  
 Db 1869 AGTAACCTTATATATTTTGGAGATTTTGGAGAAACAAAAAATTTTAAAAA 1919

## RESULT 3

ED127520 1860 bp DNA linear PAT 18-SEP-2002

LOCUS ED127520 1860 bp DNA linear PAT 18-SEP-2002

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION ED127520.1 GI:23222465

VERSION ED127520.1 JF 2002017375-A/2951

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,

Kogawa, H., Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and

Primer for synthesizing full-length cDNA and use thereof

HELIIX RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JF 2002017375-A/2951

PD 22-JAN-2002

PF 07-JUL-2000 JF 2000253172

PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO  
 PI ISHII,  
 PI YURI KAWAI, AI WAKAMATSU, TOKOYASU SUGIYAMA, KEIICHI NAGAI, PI  
 SHINICHI KOJIMA,  
 PI TETSUJI OTSUKI, HISASHI KOGA  
 PC  
 C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
 10,  
 PC C12P21/02, C1201/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
 Primer for synthesizing full-length cDNA and use thereof PH Key  
 Location/Qualifiers  
 FT CDS (129).. (1544).  
 source  
 1. 1860  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 515 a 423 c 451 g 471 t  
 ORIGIN

Query Match 98.3%; Score 1820.4; DB 6; Length 1860;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 1824; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Oy 1 GGCTAGGCGGAGAGCTTATGCTGAGAGCGGCTCGCTCCGCGCTCAGAGCGGCTTA 60  
 Db 31 GGCTAGGCGGAGAGCTTATGCTGAGAGCGGCTCGCTCCGCGCTCAGAGCGGCTTA 90  
 Oy 61 TCAGATATCTTAAAGAAAGAAACCACTGAGAAAAAATGAAATTCCTTATCTTCCAT 120  
 Db 91 TCAGATATCTTAAAGAAAGAAACCACTGAGAAAAAATGAAATTCCTTATCTTCCAT 150  
 Oy 121 TTTTGGGAGGTTCACTTTTATCCCTGCTGAGAAAGCTATGCAAGATGAGCA 180  
 Db 151 TTTTGGGAGGTTCACTTTTATCCCTGCTGAGAAAGCTATGCAAGATGAGCA 210  
 Oy 181 TCTCTAAGAGACTTTTGAAGAAATGAAGAAATAGCAGCTGTGAGATGTTCTA 240  
 Db 211 TCTCTAAGAGACTTTTGAAGAAATGAAGAAATAGCAGCTGTGAGATGTTCTA 270  
 Oy 241 AAGCAATCATCACTGCTGTTTATGTAAGCCCAAGACAGATCTTATGAGCGATTGG 300  
 Db 271 AAGCAATCATCACTGCTGTTTATGTAAGCCCAAGACAGATCTTATGAGCGATTGG 330  
 Oy 301 CACTTCTGTTGATGCTGTTGAGCCGAGCTGAGTGGCTCCAGAACTTGAAGAAAGCCA 360  
 Db 331 CACTTCTGTTGATGCTGTTGAGCCGAGCTGAGTGGCTCCAGAACTTGAAGAAAGCCA 390  
 Oy 361 TCCAAATTAATGACCAAAACCTGAGAGATGGCTGAGAAAGTTTCACTGAGGCGAG 420  
 Db 391 TCCAAATTAATGACCAAAACCTGAGAGATGGCTGAGAAAGTTTCACTGAGGCGAG 450  
 Oy 421 TGAGATATCCCACTGAGAGAGGGAGAGAAATCACTGTGATCTGAGAGCAAGATTTC 480  
 Db 451 TGAGATATCCCACTGAGAGAGGGAGAGAAATCACTGTGATCTGAGAGCAAGATTTC 510  
 Oy 481 ATAGATAGCATCTGAGGCTGTTGAGAGCAAGTGGGACCTCCAGAAAGGCAATTACAG 540  
 Db 511 ATAGATAGCATCTGAGGCTGTTGAGAGCAAGTGGGACCTCCAGAAAGGCAATTACAG 570  
 Oy 541 CAGAGTTCTGAGTGAAGCTCTTTCATGAACTGAGAGAGGCGCTCAGAAAGCAAGAG 600  
 Db 571 CAGAGTTCTGAGTGAAGCTCTTTCATGAACTGAGAGAGGCGCTCAGAAAGCAAGAG 630  
 Oy 601 GGAAGATGTTGTTTATTAACCACTTATCAATCACTAAGAGCGGTCAATACCGAA 660  
 Db 631 GGAAGATGTTGTTTATTAACCACTTATCAATCACTAAGAGCGGTCAATACCGAA 690  
 Oy 661 CGAGAGGCGGAGAGAGCTGCAAGAGTGGGCGTTTGGATCTCTATTCGATCCGCTGG 720  
 Db 691 CGAGAGGCGGAGAGAGCTGCAAGAGTGGGCGTTTGGATCTCTATTCGATCCGCTGG 750  
 Oy 721 CCTCTCTCCATCTACAGTCTCTACACAGGTAATTCAGGAATACAGAGATGCGTGCCTCA 780

```

Db      751 CCTCTTCTCCATCTACAGTCTCTCACAGGTAATCAGAAATACAGAAATGCGCTGCCA 810
Qy      781 AANFTCCAAAGCCTGATTAACGGTGGAAATGCAAAATGCTCAAGATGCGCTCTC 840
Db      811 AGATTCACAAAGCCTGATTAACGGTGGAAATGCAAAATGCTCAAGATGCGCTCTC 870
Qy      841 ATGGGATCAAAATGCTGATTCAGTAAAGATGGGGGCAAGACCTTACCAGATCTGATT 900
Db      871 ATGGGATCAAAATGCTGATTCAGTAAAGATGGGGGCAAGACCTTACCAGATCTGATT 930
Qy      901 CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAACAGGTTGATCTGTCA 960
Db      931 CTTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAACAGGTTGATCTGTCA 990
Qy      961 GTGACACTTGGACAGCTGGAGATTTGGGCGAGGTCGATGATGGCGGTGGACCT 1020
Db      991 GTGACACTTGGACAGCTGGAGATTTGGGCGAGGTCGATGATGGCGGTGGACCT 1050
Qy      1021 TTATATCATGGGAGACACTCTCATTTATTAAGATCTTGGGCTGCTCCAAAGAGACTC 1080
Db      1051 TTATATCATGGGAGACACTCTCATTTATTAAGATCTTGGGCTGCTCCAAAGAGACTC 1110
Qy      1081 TGGCGCTGCTGCTTGGACTGACAGAGAAACAAGGTGAGTGTGCTTCCAGATTATC 1140
Db      1111 TGGCGCTGCTGCTTGGACTGACAGAGAAACAAGGTGAGTGTGCTTCCAGATTATC 1170
Qy      1141 AGTTACCAAGGTAAATTTTCCACTACAGTCTGTGATGAGATCTGACGAGAACT 1200
Db      1171 AGTTACCAAGGTAAATTTTCCACTACAGTCTGTGATGAGATCTGACGAGAACT 1230
Qy      1201 TCTTACCACTGGGCTGCAATTCAGTGGCAGTGAAGAGGCGGACCATCATGAGAGG 1260
Db      1231 TCTTACCACTGGGCTGCAATTCAGTGGCAGTGAAGAGGCGGACCATCATGAGAGG 1290
Qy      1261 TTATGACCTGCTGCTGAGCCCTCAATATCACTCAAGTCTTGGAGCATGAGAGAGG 1320
Db      1291 TTATGACCTGCTGCTGAGCCCTCAATATCACTCAAGTCTTGGAGCATGAGAGAGG 1350
Qy      1321 ACATCAACTTTTGGATCCAGACTGAGTGGTGGAGGCAAGTCTACTTGTATGATCA 1380
Db      1351 ACATCAACTTTTGGATCCAGACTGAGTGGTGGAGGCAAGTCTACTTGTATGATCA 1410
Qy      1381 AGTATTTCTTTCATCATCTCCACAGGAGACACATGACTGATGGATCCAAAGCAG 1440
Db      1411 AGTATTTCTTTCATCATCTCCACAGGAGACACATGACTGATGGATCCAAAGCAG 1470
Qy      1441 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGACATGGAAGAA 1500
Db      1471 TGAATGTTGCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTCAGACATGGAAGAA 1530
Qy      1501 TGTGCTCCTAGGTCTAGAAACAGTAAAGAAAGAAAGCTTTCACTGCTTCCAGAAATC 1560
Db      1531 TGTGCTCCTAGGTCTAGAAACAGTAAAGAAAGAAAGCTTTCACTGCTTCCAGAAATC 1590
Qy      1561 CTGGGCTGCAACTTTGGAAACCTCTCTTCAATTAATTAATTAATTAATTAATTAAT 1620
Db      1591 CTGGGCTGCAACTTTGGAAACCTCTCTTCAATTAATTAATTAATTAATTAATTAAT 1650
Qy      1621 AAGCAACACTTATTTCACTGCTTCTGTTATTAATTTCTTGTATGATCACTTCAAAATCT 1680
Db      1651 AAGCAACACTTATTTCACTGCTTCTGTTATTAATTTCTTGTATGATCACTTCAAAATCT 1710
Qy      1681 GATTCTAGAAAAAGAAATCAATTCCTCCCTCCCTCCACCAATGATCAACATATGATA 1740
Db      1711 GATTCTAGAAAAAGAAATCAATTCCTCCCTCCCTCCACCAATGATCAACATATGATA 1770
Qy      1741 GGGATTAAGTGGGGGCACTTTCTTATATCACTTAAATCACTTAAATCACTTAAATCA 1800
Db      1771 GGGATTAAGTGGGGGCACTTTCTTATATCACTTAAATCACTTAAATCACTTAAATCA 1830
Qy      1801 AGTAAACACTTAATTAATTTTGGAAAGATC 1830

```

```

Db      1831 AGTAAACACTTAATTAATTTTGGAAAGATC 1860
RESULT 4
AK075132
LOCUS
DEFINITION
AK075132 1860 bp mRNA linear PRI 03-SEP-2002
Homo sapiens cDNA FLJ90651 fis, clone PLACE1004482, moderately
similar to Rattus norvegicus hematopoietic lineage switch 2 related
protein (H182-tp) mRNA.
ACCESSION
AK075132
VERSION
AK075132.1 GI:22761022
KEYWORDS
Oligo capping, fis (full insert sequence).
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
Isogai,T., Ota,T., Nishikawa,T., Hayaashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Satou,K., Yamamoto,U., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahari,K., Masubo,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1860)
Isogai,T. and Otsuki,T.
Direct Submission (25-MAR-2002) Takao Isogai, Helix Research Institute,
Submitted (25-MAR-2002) 1532-3 Yama, Kisarazu, Chiba 292-0812, Japan
Genomics Laboratory; 1532-3 Yama, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction;
Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection; Helix Research Institute (supported
by Japan Key Technology Center etc.).
FEATURES
source
1. 1860
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1004482"
/tissue_type="placenta"
/clone_lib="PLACE1"
/notes="Cloning vector: pME18SFL3"
129..1547
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAC11423.1"
/db_xref="GI:22761023"
/translation="MKRLIFAFRGVHLISLCSGKAI CKNGISKRTFEEIEEIASCG
DVKKAILINLAVYKQAKQNSYERLALVDYTPGLSGKNLEKAIQIYQWLQODGLEK
VHLEPVRI PHMERGESAVMLEPRIRHIALIGSSSIGTPEGITAEVLVTSFDEIQ
RRSPBARGKIVVNOPYNSRTVOYRTQGVAAKVGALASLIRSVASISYSPHQ
IOXYOCVPRKIPTRACITVDAEMSMASGIVTIOIKGATVPEPTDSFNTVARTT
GSKYPRQVVLVSGHLSWVDVGAGMDDGGAFISWEALSLIKDIGLRPKRTLRVLT
AEQGGVGAFOYVQLHKNVINSLSLVNESDAGTFLPTGLQFTSSEKRAIAVEEVSIL
QPNLTQVLSHGGSTIDINFWIOAGVPASILLDDLYKFFPHSHSGDMTWDPKQNV
AAAVMAVSVYVADMEEMLPRS"
BASE COUNT
515 a 423 c 451 g 471 t
ORIGIN
Query Match 98.3%; Score 1820.4; DB 9; Length 1860;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1824; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy      1 GCGTAGGCGCGGAGCTTAAAGTCCGAGGAGCGCGCTCGCGCGGCTCAGAGCGCGCTTA 60
Db      31 GCGTAGGCGCGGAGCTTAAAGTCCGAGGAGCGCGCTCGCGCGGCTCAGAGCGCGCTTA 90
Qy      61 TCAGATTATCTTAACAAGAAAAACCAACTGGAAGAAAAAATGAAATCTTATCTTCGAT 120

```

```

Db      91 TCAGATATCTTAACAGAAAACCAACTGGAAAAAAATGAATTCTTATCTTCGAT 150
Qy      121 TTTTGGGAGGTGTACCTTTATCCCTGTCTGGGAAACCTATATGCAAGATGGCA 180
Db      151 TTTTGGGAGGTGTACCTTTATCCCTGTCTGGGAAACCTATATGCAAGATGGCA 210
Qy      181 TCTTAAGAGCACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGGAGATGTGGTA 240
Db      211 TCTTAAGAGCACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGGAGATGTGGTA 270
Qy      241 AAGCAATCATCAACTAGCTGTATATGTAAAGCCAGAAACAGATCTATAGCGATTGG 300
Db      271 AAGCAATCATCAACTAGCTGTATATGTAAAGCCAGAAACAGATCTATAGCGATTGG 330
Qy      301 CACTCTGGTGTATCTGTGGAGCCAGCTGAGTGGCTCCAGAAACCTAGAAAGCA 360
Db      331 CACTCTGGTGTATCTGTGGAGCCAGCTGAGTGGCTCCAGAAACCTAGAAAGCA 390
Qy      361 TCCAAATTATGTACCAAAAACCTGAGCAAGATGGGCTGGAGAAAGTTCACTGGAGCCAG 420
Db      391 TCCAAATTATGTACCAAAAACCTGAGCAAGATGGGCTGGAGAAAGTTCACTGGAGCCAG 450
Qy      421 TGAGAAATACCCCACTGGAGAGAGGAGAAATACAGCTGTGTATGTGTAGCCAGAAATTC 480
Db      451 TGAGAAATACCCCACTGGAGAGAGGAGAAATACAGCTGTGTATGTGTAGCCAGAAATTC 510
Qy      481 ATAAATAGCCATCTCTGGGTCTTTGGAGAGATGGGACTCTCCAGAAAGCATTTACAG 540
Db      511 ATAAATAGCCATCTCTGGGTCTTTGGAGAGATGGGACTCTCCAGAAAGCATTTACAG 570
Qy      541 CAGAAGTCTGGTGTGACCTCTTGTGATGATGACAGAGAGGCTCAGAGAGCAAG 600
Db      571 CAGAAGTCTGGTGTGACCTCTTGTGATGATGACAGAGAGGCTCAGAGAGCAAG 630
Qy      601 GGAAGATTGTGTTTATTAACAACCTTACATCACTCAAGAGCGGTGCATAACCGAA 660
Db      631 GGAAGATTGTGTTTATTAACAACCTTACATCACTCAAGAGCGGTGCATAACCGAA 690
Qy      661 CGCAGAGGCGGTGGAGAGCTGCAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGG 720
Db      691 CGCAGAGGCGGTGGAGAGCTGCAAGGTGGGGCTTTGGCATCTCTCATTCGATCCGTGG 750
Qy      721 CCTCCTTCTCATCTACAGTCTCTCAACAGATTAACAGAAATGATGCAAGATGCGCTTCTC 780
Db      751 CCTCCTTCTCATCTACAGTCTCTCAACAGATTAACAGAAATGATGCAAGATGCGCTTCTC 810
Qy      781 AAATTCACAGAGCTGTATTTACGGTGAAGATGCAAGAAATGATGCAAGATGCGCTTCTC 840
Db      811 AGATTCACAGAGCTGTATTTACGGTGAAGATGCAAGAAATGATGCAAGATGCGCTTCTC 870
Qy      841 ATGGATCAAAAATTTGTCAATTCAGCTTAAGATGGGGGCAAAAGCTTAACCAATGATGAT 900
Db      871 ATGGATCAAAAATTTGTCAATTCAGCTTAAGATGGGGGCAAAAGCTTAACCAATGATGAT 930
Qy      901 CTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAAGAGTTGACTGTGCA 960
Db      931 CTTCAACACTGTAGCAGAGATCACTGGAGCAAAATTCAGAAAGAGTTGACTGTGCA 990
Qy      961 GTGACATCTGGAGAGCTGGAGATGTTGGGAGAGGTCAGATGATGAGCGGTGGAGCT 1020
Db      991 GTGACATCTGGAGAGCTGGAGATGTTGGGAGAGGTCAGATGATGAGCGGTGGAGCT 1050
Qy      1021 TTATATCATGGAGAGCACTTCACTTAATTAAGATCTTGGGCTGGCTCCAAAGAGACTC 1080
Db      1051 TTATATCATGGAGAGCACTTCACTTAATTAAGATCTTGGGCTGGCTCCAAAGAGACTC 1110
Qy      1081 TGCGGCTGGTGTCTGAGCTGCAAGAAACAAGGTGAGTGGTCTTCCAGATTAATC 1140
Db      1111 TGCGGCTGGTGTCTGAGCTGCAAGAAACAAGGTGAGTGGTCTTCCAGATTAATC 1170
Qy      1141 AGTTACACAGAGTAATATTTTCAACTACAGCTGCTGTGATGAGAGTCTGACGAGAACT 1200

```

```

Db      1171 AGTTACACAGAGTAATATTTTCCAACTACAGTCTGTGTATGAGTCTGACGCAAGAACT 1230
Qy      1201 TCTTACCACACTGGGCTGCAATTTCACTGGCAGGTGAAGAGCCAGGCGCATCATGAGAGAG 1260
Db      1231 TCTTACCACACTGGGCTGCAATTTCACTGGCAGGTGAAGAGCCAGGCGCATCATGAGAGAG 1290
Qy      1261 TTATAGAGCTGCTCAGCCCTCAATATCACTCAGGTCTGAGCCATGAGAGAGAGAGCAG 1320
Db      1291 TTATAGAGCTGCTCAGCCCTCAATATCACTCAGGTCTGAGCCATGAGAGAGAGAGCAG 1350
Qy      1321 ACATCAACTTTTGGATTCAGAGCTGAGTGGCTGAGAGCCAGTCTTCTGATGATTAATCA 1380
Db      1351 ACATCAACTTTTGGATTCAGAGCTGAGTGGCTGAGAGCCAGTCTTCTGATGATTAATCA 1410
Qy      1381 AGTATTTCTTCTTCATCACTCCAGGAGAGACCATGACTGTATGATGATCAAGAGCA 1440
Db      1411 AGTATTTCTTCTTCATCACTCCAGGAGAGACCATGACTGTATGATGATCAAGAGCA 1470
Qy      1441 TGAATGTGCTGCTGCTGTTTGGGCTGTTTCTTATATGTTGTGACAGACATGAGAGAA 1500
Db      1471 TGAATGTGCTGCTGCTGTTTGGGCTGTTTCTTATATGTTGTGACAGACATGAGAGAA 1530
Qy      1501 TGCTGCTAGGCTCTAGAAACAGTAAGAAAGAAAGCTTTTCTATGCTTGGCCAGAAATC 1560
Db      1531 TGCTGCTAGGCTCTAGAAACAGTAAGAAAGAAAGCTTTTCTATGCTTGGCCAGAAATC 1590
Qy      1561 CTGGGTCTGCAACTTTGAGAAACCTCTTCACTACATTAATTCATTCATTCATTCATTC 1620
Db      1591 CTGGGTCTGCAACTTTGAGAAACCTCTTCACTACATTAATTCATTCATTCATTCATTC 1650
Qy      1621 AAGCAACACTATTTATGATGCTTCTGTATATATCTTCTGTGATGATCTTCCAAATTCCT 1680
Db      1651 AAGCAACACTATTTATGATGCTTCTGTATATATCTTCTGTGATGATCTTCCAAATTCCT 1710
Qy      1681 GATTCAGAAAAAGAAATCATTTCTCCCTCCCTCCACACATGATGATCAATATGATGTA 1740
Db      1711 GATTCAGAAAAAGAAATCATTTCTCCCTCCCTCCACACATGATGATCAATATGATGTA 1770
Qy      1741 GGGATTCAGTGGGGGGAATTTCTTATATGACCTTAAATAAGCTGTTTCCACTTTAA 1800
Db      1771 GGGATTCAGTGGGGGGAATTTCTTATATGACCTTAAATAAGCTGTTTCCACTTTAA 1830
Qy      1801 AGTAACACTTAATTAATTTTGGAGATC 1830
Db      1831 AGTAACACTTAATTAATTTTGGAGATC 1860

RESULT 5
AX006440
LOCUS      AX006440      1860 bp      DNA      linear      PAT 06-SEP-2000
DEFINITION Sequence 5 from Patent WO0004157.
ACCESSION  AX006440
VERSION     AX006440.1  GI:9994575
KEYWORDS
SOURCE
ORGANISM   Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
            Neefs,J.M., Peeters,D.C. and Pangalos,M.
            Cloning and characterisation of novel mammalian peptidases
            Patent: WO 0004157-A 5 27-JAN-2000;
            JANSSEN PHARMACEUTICA NV (BE); NEEFS JEAN MARC EDMOND FERNAND (BE);
            PEETERS DANIELLE CELINE GEORGE (BE); PANGALOS MENILAS (GB)
FEATURES
            source
            1. 1860
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT      518 a      420 c      451 g      471 t
ORIGIN
Query Match      96.3%; Score 1783.2; DB 6; Length 1860;

```

Best Local Similarity 99.3%; Pred. No. 0; Matches 1791; Conservative 0; Mismatches 13; Indels 0; Gaps 0;			
Qy	31	GCCTCCGTCGCGCGGTCAGAGCGCCCTATCAGATTAATCTTAACAAGAAAACCAACTGG	90
Db	57	GCGCGAGCTTAGTTCCTGGAGCGCCCTATCAGATTAATCTTAACAAGAAAACCAACTGG	116
Qy	91	AAAAAAAAAATGAATTCCTTATCTTCCGATTTTCCGAGTGTTCACCTTTATCCCTGT	150
Db	117	AAAAAAAAAATGAATTCCTTATCTTCCGATTTTCCGAGTGTTCACCTTTATCCCTGT	176
Qy	151	GCTCTGGAAAAGCTATATGCAAGATGGCATCTCTAAGAGACTTTTGAAGAAATTAAG	210
Db	177	GCTCTGGAAAAGCTATATGCAAGATGGCATCTCTAAGAGACTTTTGAAGAAATTAAG	236
Qy	211	AAGAAATAGCGAGCTGAGATGTTGCTAAGCAATCACTAGCTGTTATATGTA	270
Db	237	AAGAAATAGCGAGCTGAGATGTTGCTAAGCAATCACTAGCTGTTATATGTA	296
Qy	271	AAGCCAGAAACAGATCCATATGAGCGATTTGCACTTCTGTTGATAGTGGACCCAGAC	330
Db	297	AAGCCAGAAACAGATCCATATGAGCGATTTGCACTTCTGTTGATAGTGGACCCAGAC	356
Qy	331	TGAGTGGCTCCAGAACCTAGAAAAGCCATCCAATTTATGTACAAAACCTGACGAC	390
Db	357	TGAGTGGCTCCAGAACCTAGAAAAGCCATCCAATTTATGTACAAAACCTGACGAC	416
Qy	391	ATGGGCTGAGAAAGTTCACTGAGACGACGAGAAATCCCACTGGAGAGGGAGAG	450
Db	417	ATGGGCTGAGAAAGTTCACTGAGACGACGAGAAATCCCACTGGAGAGGGAGAG	476
Qy	451	AATCAGCTGTGATCTGAGAGCCAGAAATTCATTAAGATAGCCATCTGGGTTTGGACGA	510
Db	477	AATCAGCTGTGATCTGAGAGCCAGAAATTCATTAAGATAGCCATCTGGGTTTGGACGA	536
Qy	511	GCATTTGGAATCTCTCAAGAGGCAATTAACAGCAAGTTCTGGTGTGACCTCTTTCATG	570
Db	537	GCATTTGGAATCTCTCAAGAGGCAATTAACAGCAAGTTCTGGTGTGACCTCTTTCATG	596
Qy	571	AACATGCAAGAAAGGCTCAGAGCAAGAGGAAAGATTGTTTATTAACAACCTTACA	630
Db	597	AACATGCAAGAAAGGCTCAGAGCAAGAGGAAAGATTGTTTATTAACAACCTTACA	656
Qy	631	TCAACTCTCAAGAGCGGTGCAATACGGAACGAGGGGGCGGTGGAAGCTGCCAAGGTG	690
Db	657	TCAACTCTCAAGAGCGGTGCAATACGGAACGAGGGGGCGGTGGAAGCTGCCAAGGTG	716
Qy	691	GGGCTTTGGCATCTCTGATTCGATCCGTCCTCTCTCCATCTACAGTCTCACAAG	750
Db	717	GGGCTTTGGCATCTCTGATTCGATCCGTCCTCTCTCCATCTACAGTCTCACAAG	776
Qy	751	GTATTCAGGAATACCAAGATGGCGTCCCAAAATTCACAACGCTGTATTAACGCTGAAG	810
Db	777	GTATTCAGGAATACCAAGATGGCGTCCCAAAATTCACAACGCTGTATTAACGCTGAAG	836
Qy	811	ATGCAGAAATGATGTCAAGATGCTCTCATGGGATTAATAATGTACCTAAGCTAAAG	870
Db	837	ATGCAGAAATGATGTCAAGATGCTCTCATGGGATTAATAATGTACCTAAGCTAAAG	896
Qy	871	TGGGGGCAAAAGACCTACCAAGATCTGATTCCTTCAACACTGTAGCAGAGATCACTGGGA	930
Db	897	TGGGGGCAAAAGACCTACCAAGATCTGATTCCTTCAACACTGTAGCAGAGATCACTGGGA	956
Qy	931	GCAAAATATCCAGAACAGTTTACTGTGATGACATCTGGAACAAGCTGGAGATTTGGGC	990
Db	957	GCAAAATATCCAGAACAGTTTACTGTGATGACATCTGGAACAAGCTGGAGATTTGGGC	1016
Qy	991	AGGATGATGATGATGATGGGGGCTTATATATATGAGGAGCACTCTCACTTATTA	1050
Db	1017	AGGATGATGATGATGATGGGGGCTTATATATATGAGGAGCACTCTCACTTATTA	1076
Qy	1051	AAGATCTTGGGCTGCTCAAAAGAGACTCTGCGGCTGTGCTCTGCACTGCAGAGAAC	1110

Db	1077	AAGATCTTGGGCTGCTCCAAAGAGACTCTGCGGCTGTGCTCTGCACTGCAGAGAAC	1136
Qy	1111	AAGTGGAGTTGGTGGCTTCAGATTAATCAGTATACCAAGGTAATTAATTCACACTACA	1170
Db	1137	AAGTGGAGTTGGTGGCTTCAGATTAATCAGTATACCAAGGTAATTAATTCACACTACA	1196
Qy	1171	GTCGTGATGAGTGTGAGCAGCAAGAACTTTCTTACCACCTGGGCTCAATTCATGGCA	1230
Db	1197	GTCGTGATGAGTGTGAGCAGCAAGAACTTTCTTACCACCTGGGCTCAATTCATGGCA	1256
Qy	1231	GTCGAGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG	1290
Db	1257	GTCGAGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG	1316
Qy	1291	CTCAGTCTGAGGCAATGAGAGGAGCAGACATCACTTTTGAATCCAGCTGGATGC	1350
Db	1317	CTCAGTCTGAGGCAATGAGAGGAGCAGACATCACTTTTGAATCCAGCTGGATGC	1376
Qy	1351	CTGAGGCACTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1410
Db	1377	CTGAGGCACTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1436
Qy	1411	ACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1470
Db	1437	ACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1496
Qy	1471	TTTCTTATGTTGTTGCAAGATGAGAAATGCTGCTAGTCTGCTAGTCTGCTAGTCTG	1530
Db	1497	TTTCTTATGTTGTTGCAAGATGAGAAATGCTGCTAGTCTGCTAGTCTGCTAGTCTG	1556
Qy	1531	GAAGCTTTTCACTGCTCTGAGGAGAACTCTGGGCTGTGCAACTTTGGAAGAACTCTT	1590
Db	1557	GAAGCTTTTCACTGCTCTGAGGAGAACTCTGGGCTGTGCAACTTTGGAAGAACTCTT	1616
Qy	1591	CACATTAACAATTCATCAATTCATCTTCAAGACAACTCTATTCATGCTTCTGTTA	1650
Db	1617	CACATTAACAATTCATCAATTCATCTTCAAGACAACTCTATTCATGCTTCTGTTA	1676
Qy	1651	TTATCTTTCTTGATCTTTTCCAAATTTCTGATTTCTGAAAGAAAGATCATTCCTCCTC	1710
Db	1677	TTATCTTTCTTGATCTTTTCCAAATTTCTGATTTCTGAAAGAAAGATCATTCCTCCTC	1736
Qy	1711	CCTCCACACATAGAAATCAACATATGATGAGGATTAACAGTGGGGGCAATTCCTTATATC	1770
Db	1737	CCTCCACACATAGAAATCAACATATGATGAGGATTAACAGTGGGGGCAATTCCTTATATC	1796
Qy	1771	ACCTCTTAAAGAAATGTTTCCACTTTTAAAGTAACAATTAATTTTGGAGATC	1830
Db	1797	ACCTCTTAAAGAAATGTTTCCACTTTTAAAGTAACAATTAATTTTGGAGATC	1856
Qy	1831	TTCTG 1834	
Db	1857	TTCTG 1860	
RESULT 6			
AF119386			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
N-acetyl-aspartyl-alpha-glutamate			

JOURNAL  
J. Biol. Chem. 274 (17), 11742-11750 (1999)  
MEDLINE  
9923495  
PUBMED  
10206990  
2 (bases 1 to 1794)  
Gingras, R., Richard, C., El-Alfy, M., Morales, C.R., Potier, M. and  
Pehzetsky, A.V.  
AUTHORS  
TITLE  
Direct Submission  
Submitted (12-JUN-1999) Medical Genetics, Sainte-Justine Hospital,  
JOURNAL  
Montreal University, 3175 Cote Sainte-Catherine, Montreal, QH H3T  
1C5, Canada

# FEATURES

## source

Location/Qualifiers  
1. .1794  
/organism="Homo sapiens"  
/mol\_type="rRNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8q22.2; between D8S257 and D8S270"  
1. .1794  
/gene="PGCP"  
25. .1650  
/gene="PGCP"  
/EC\_number="3.4.17.21"  
/function="endopeptidase"  
/note="member of M28 peptidase family; similar to  
glutamate carboxypeptidase II and to  
N-acetyl-aspartyl-alpha-glutamate carboxypeptidase  
(NALADase); similar to sequences with GenBank Accession  
Numbers R18560 and G07404; prostate-specific membrane  
antigen (PSMA)"  
/codon\_start=1  
/product="blood plasma glutamate carboxypeptidase  
precursor"  
/protein\_id="AAD31418.1"  
/db\_xref="GI:4877698"  
/translation="MKFLIFAFPGVHLISCSGKAIKNGISKRTPEEIKELIASCS  
DVAKILAVYKQNRYSERIALLVDTGPRLSGSKLEKAIQIMQNOQGLEK  
VHLBPVRI PHMERGESAVMLPRRIKIAIIGLSISGTPPEGITAEVLVYTSDEILQ  
RASEARGIYVNPYINVSRTVYRQGAIVEAKVAKALASLIRSVASFEIYSPHTS  
IOETODGVPKIPFACITVEDABMSRMKSHGKIYIQLMGAKTYPDTSFNTVAEIT  
GSKYPEQVIVLSGLDSWDVQAGADDGGAFLISWEALSLIKDGLRKRFLRLVLT  
AEEQGVGAFOYQLAKVNIISYSLVMSDGTFLPTGLQTSSEKARAIIEVWSIL  
QPLNTVYLSHGEGDINFMIOAGVPGASLIDDLKYFFPHSHSDTWTMDPSMML  
LLRFLFLMLLQTKKCCIGPRNKKETFCFMPGIGLQIMTKPTLHITISSNSSK  
HNSISCELLSFLISKSDSKRHSFLPPTT"  
25. .96  
/gene="PGCP"  
157. .1647  
/gene="PGCP"  
/product="blood plasma glutamate carboxypeptidase"  
205. .207  
/gene="PGCP"  
/note="putative; glycosylation site"  
559. .561  
/gene="PGCP"  
/note="putative; glycosylation site"  
1081. .1083  
/gene="PGCP"  
/note="putative; glycosylation site"  
1090. .1092  
/gene="PGCP"  
/note="putative; glycosylation site"  
1210. .1212  
/gene="PGCP"  
/note="putative; glycosylation site"

## gene

## CDS

BASE COUNT 522 a 422 g 457 t 2 others  
ORIGIN

Query Match 95.0%; Score 1758.6; DB 9; Length 1794;  
Beet Local Similarity 99.7%; Pred.No. 0;  
Matches 1772; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

75 CAAGAAACCAACTGGAAAGAAATGAATTCCTTATCTTGCAATTTTTCGGTGTGT 134.

|||||  
Db 1 CAAGAAACCAACTGGAAAGAAATGAATTCCTTATCTTGCAATTTTTCGGTGTGT 60  
Oy 135 CACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGATGGCATCTCTAAGAGACT 194  
Db 61 CACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGATGGCATCTCTAAGAGACT 120  
Oy 195 TTGAGAAATTAAGAAAGAAATATGACAGCTGTGGAGATGTGGTAAAGCAATCATCAAC 254  
Db 121 TTTAAGAAATTAAGAAAGAAATATGACAGCTGTGGAGATGTGGTAAAGCAATCATCAAC 180  
Oy 255 CTACCTGTTATATGTTAAAGCCAGAAAGATCTATAGAGCAATTTGGCACTTCTGTTGAT 314  
Db 181 CTACCTGTTATATGTTAAAGCCAGAAAGATCTATAGAGCAATTTGGCACTTCTGTTGAT 240  
Oy 315 ACTGTTGAGCCCAAGCTGAGTGTCTCAAGAACTTAAGAAAGCAATTTATATGATAC 374  
Db 241 ACTGTTGAGCCCAAGCTGAGTGTCTCAAGAACTTAAGAAAGCAATTTATATGATAC 300  
Oy 375 CAAGAACTGAGCAAGATGGGCTGGAGAAAGTTCACTGGAGCCAGTGAATACCCGAC 434  
Db 301 CAAGAACTGAGCAAGATGGGCTGGAGAAAGTTCACTGGAGCCAGTGAATACCCGAC 360  
Oy 435 TGGAGAGGGGAGAAAGATCAGCTGTATGCTGGAGCCAGAAATTCATTAAGATAGCATC 494  
Db 361 TGGAGAGGGGAGAAAGATCAGCTGTATGCTGGAGCCAGAAATTCATTAAGATAGCATC 420  
Oy 495 CTGGGCTCTTGACAGACATTTGGGATCTCTCCAGAAAGCATTAACAGCAAGTTCTGTG 554  
Db 421 CTGGGCTCTTGACAGACATTTGGGATCTCTCCAGAAAGCATTAACAGCAAGTTCTGTG 480  
Oy 555 GTGACCTCTTTCATGTAAGTGCAGAGAGGGCCCTCAGAGCAAGAGGAAAGATGTTGTT 614  
Db 481 GTGACCTCTTTCATGTAAGTGCAGAGAGGGCCCTCAGAGCAAGAGGAAAGATGTTGTT 540  
Oy 615 TATAACCAACCTTATCATCACTAATCAAGACGCTGCAATACCAAGCGAGGGGCGGATG 674  
Db 541 TATAACCAACCTTATCATCACTAATCAAGACGCTGCAATACCAAGCGAGGGGCGGATG 600  
Oy 675 GAAGCTGCAAGGTTGGGGGCTTTGGCATCTCTCATTTGATTCGTGCGCTCTCTTCATC 734  
Db 601 GAAGCTGCAAGGTTGGGGGCTTTGGCATCTCTCATTTGATTCGTGCGCTCTCTTCATC 660  
Oy 735 TAGAGTCTTCAACAGGATTAATCAGAAATACAGAGATGGCGGCCAAATTCGAAGGCC 794  
Db 661 TAGAGTCTTCAACAGGATTAATCAGAAATACAGAGATGGCGGCCAAATTCGAAGGCC 720  
Oy 795 TGTATTAAGTGAAGATGCAAGAAATGATCAAGAAATGGCTTCTCATGGGATCAAAAT 854  
Db 721 TGTATTAAGTGAAGATGCAAGAAATGATCAAGAAATGGCTTCTCATGGGATCAAAAT 780  
Oy 855 GTCATTTAGCTTAAGATGAGGGGCAAGAACTTAACCAATCTGATTTCTTCAACACTGA 914  
Db 781 GTCATTTAGCTTAAGATGAGGGGCAAGAACTTAACCAATCTGATTTCTTCAACACTGA 840  
Oy 915 GCAGAGATCACTGGAGCAAAATATCAAGAAAGGTTGATCTGACAGTGAAGCATCTGAG 974  
Db 841 GCAGAGATCACTGGAGCAAAATATCAAGAAAGGTTGATCTGACAGTGAAGCATCTGAG 900  
Oy 975 AGCTGGATGTTGGGAGAGGTTGCCATGATGATGCGGTGAGGCTTTATATCAATGGGAA 1034  
Db 901 AGCTGGATGTTGGGAGAGGTTGCCATGATGATGCGGTGAGGCTTTATATCAATGGGAA 960  
Oy 1035 GCACCTTCATTTAAGATCTTGGGCTGCTCAAGAAAGAGACTTGGCGGCTGTGCTC 1094  
Db 961 GCACCTTCATTTAAGATCTTGGGCTGCTCAAGAAAGAGACTTGGCGGCTGTGCTC 1020  
Oy 1095 TGAGCTGCAAGAAAGAAAGGTTGAGTGTGGCTTCAAGATTAATCAAGGTA 1154  
Db 1021 TGAGCTGCAAGAAAGAAAGGTTGAGTGTGGCTTCAAGATTAATCAAGGTA 1080  
Oy 1155 AATATTTCAACTACAGTCTGTGATGAGATCTGACGAGAACTTTTACCACTGGG 1214

Db 1081 AATATTCACATCACTGCTGATGAGTGTGACGAGAACCTTCTTACCACTGGG 1140  
Qy 1215 CTGCAATTCAGTGGCAGTGAAGAGCCAGGCGCATCATGAGAGGTTATAGCGCTG 1274  
Db 1141 CTGCAATTCAGTGGCAGTGAAGAGCCAGGCGCATCATGAGAGGTTATAGCGCTG 1200  
Qy 1275 CAGCCCTCAATATCACTGAGTCTTGAGCCATGAGAGGAGACATCACTTTTG 1334  
Db 1201 CAGCCCTCAATATCACTGAGTCTTGAGCCATGAGAGGAGACATCACTTTTG 1260  
Qy 1335 ATCCAACTGAGAGGCTTGAGCCAGTCTATGATGATCATATACAGATTTCTTCTC 1394  
Db 1261 ATCCAACTGAGAGGCTTGAGCCAGTCTATGATGATCATATACAGATTTCTTCTC 1320  
Qy 1395 CATCACTCCAGAGAGACCATGATGATGATGATGATGATGATGATGATGATGATG 1454  
Db 1321 CATCACTCCAGAGAGACCATGATGATGATGATGATGATGATGATGATGATGATG 1379  
Qy 1455 GCTGTTGGGCTGTTGTTCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1514  
Db 1380 GCTGTTGGGCTGTTGTTCTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1439  
Qy 1515 TAGAAACAGTGAAGAAAGAGTTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 1574  
Db 1440 TAGAAACAGTGAAGAAAGAGTTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCT 1499  
Qy 1575 TTGAAACAGTCTCTGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1634  
Db 1500 TTGAAACAGTCTCTGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1559  
Qy 1635 TTGATGCTTCTGTTATATCTTCTGTTATCTTCTGTTATCTTCTGTTATCTTCTGTT 1694  
Db 1560 TTGATGCTTCTGTTATATCTTCTGTTATCTTCTGTTATCTTCTGTTATCTTCTGTT 1619  
Qy 1695 GAATCATCTTCCCTCCCTCCCAACATAGATCAATAGTGAATTAAGTGAAGTGA 1754  
Db 1620 GAATCATCTTCCCTCCCTCCCAACATAGATCAATAGTGAATTAAGTGAAGTGA 1679  
Qy 1755 GGCAATCTTATATCACTCTTAAACATTTGTTTCACTTAAAGTAAACCTTAT 1814  
Db 1680 GGCAATCTTATATCACTCTTAAACATTTGTTTCACTTAAAGTAAACCTTAT 1739  
Qy 1815 AATTTTGGAAAGATCTCTGAAAAAATTTTAAAAA 1851  
Db 1740 AATTTTGGAAAGATCTCTGAAAAAATTTTAAAAA 1776

RESULT 7  
BD105816  
LOCUS BD105816 1778 bp DNA linear PAT 18-SEP-2002  
DEFINITION Secretory protein.  
ACCESSION BD105816  
VERSION BD105816.1 GI:23200634  
KEYWORDS JP 2002502234-A/7.  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 1778)  
AUTHORS Jacobs,K., McCoy,J.M., Racie,L.A., Lavallie,E.R., Merberg,D. and Spaulding,V.  
TITLE Secretory protein  
JOURNAL Patent: JP 2002502234-A 7 22-JAN-2002;  
GENETICS INSTITUTE INC.  
COMMENT  
PN JP 2002502234-A/7  
PF 16-APR-1997 JP 1997537384  
PR 18-APR-1996 US 08/634325, 13-JAN-1997 US 08/783520 PI  
KENNETH JACOBS, JOHN M MCCOY, LISA A RACIE, EDWARD R LAVALLIE, PI  
DAVID MERBERG,  
PI VIKKI SPAULDING  
PC C12N15/12, C07K14/47, A61K38/17  
CC Strandedness: Double;  
Topology: Linear;

FEATURES  
source 1.1778  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
BASE COUNT 514 a 386 c 422 g 456 t  
ORIGIN  
Query Match 94.6%; Score 1751; DB 6; Length 1778;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1757; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
85 AACTGAAAAAATGAATATCTTATCTTGGATTTTCGGGTGTTTACCTTTAT 144  
3 AGCTTGACAGAGATATTAATCTTATCTTGGATTTTCGGGTGTTTACCTTTAT 62  
Qy 145 CCTGTGCTCTGGGAAAGCTATATGCAAGATGGCATCTTAAGAGACTTTTGAAGAA 204  
Db 63 CCTGTGCTCTGGGAAAGCTATATGCAAGATGGCATCTTAAGAGACTTTTGAAGAA 122  
Qy 205 TAAAGAAAGAAATGCGAGCTGTGAGATGTGTGTAAGCAATATCACTAGCTTT 264  
Db 123 TAAAGAAAGAAATGCGAGCTGTGAGATGTGTGTAAGCAATATCACTAGCTTT 182  
Qy 265 ATGTAAAGCCCAAGACAGATCTATGAGGATTGGCACTTGGTGTATGTTGAGAC 324  
Db 183 ATGTAAAGCCCAAGACAGATCTATGAGGATTGGCACTTGGTGTATGTTGAGAC 242  
Qy 325 CCAAGTGAAGTGGCTCCAGAACCTAGAAAAAGCCATCAATATATGACCAAAACCTGC 384  
Db 243 CCAAGTGAAGTGGCTCCAGAACCTAGAAAAAGCCATCAATATATGACCAAAACCTGC 302  
Qy 385 AGCAAGTGGCTGTGAAGAAAGTTCACTGAGCCATGAGATACCACTGGGAGAGG 444  
Db 303 AGCAAGTGGCTGTGAAGAAAGTTCACTGAGCCAGTGAAGATACCACTGGGAGAGG 362  
Qy 445 GAGAAAGTCAAGCTGTATGTGAGCCAAAGATTTCAATGATATGCCATCTGGGTCTG 504  
Db 363 GAGAAAGTCAAGCTGTATGTGAGCCAAAGATTTCAATGATATGCCATCTGGGTCTG 422  
Qy 505 GCACAGCATTTGGGACTCTCCAGAAAGCATTTACAGAGAAAGTTCTGGTGTGA 564  
Db 423 GCACAGCATTTGGGACTCTCCAGAAAGCATTTACAGAGAAAGTTCTGGTGTGA 482  
Qy 565 TCGATGAATGTCAGAGAGAGGCTCAGAGAGCAAGAGAGAAATGTTGTTATACCAAC 624  
Db 483 TCGATGAATGTCAGAGAGAGGCTCAGAGAGCAAGAGAGAAATGTTGTTATACCAAC 542  
Qy 625 CTTCATCACTACTCAAGAGAGGTTGCAATACCAAGAGAGGAGGAGGAGGAGGAGGAGG 684  
Db 543 CTTCATCACTACTCAAGAGAGGTTGCAATACCAAGAGAGGAGGAGGAGGAGGAGGAGG 602  
Qy 685 AGGTGGGGGCTTTGGCATCTCTCATTCGATCCGTTGCTCTTCATCATCACTAGCTTC 744  
Db 603 AGGTGGGGGCTTTGGCATCTCTCATTCGATCCGTTGCTCTTCATCATCACTAGCTTC 662  
Qy 745 ACACAGGTATTCAGAAATACAGAGATGGCTGCCAAATTTCAACAGCTGTATTAAG 804  
Db 663 ACACAGGTATTCAGAAATACAGAGATGGCTGCCAAATTTCAACAGCTGTATTAAG 722  
Qy 805 TGGAAAGTGAAGAAATGATGCAAGATGCTTCTCATGGAATCAAAATTTCTCACTAGC 864  
Db 723 TGGAAAGTGAAGAAATGATGCAAGATGCTTCTCATGGAATCAAAATTTCTCACTAGC 782  
Qy 865 TAAAGATGGGGGCAAGACCTACCAAGATCTGATTTCTTCAACACTGTAGCAGATCA 924  
Db 783 TAAAGATGGGGGCAAGACCTACCAAGATCTGATTTCTTCAACACTGTAGCAGATCA 842  
Qy 925 CTGGAGCAATATCAAGAAAGGTTGATCTGCTGATGACATCTGAGACAGCTGGGATG 984  
Db 843 CTGGAGCAATATCAAGAAAGGTTGATCTGCTGATGACATCTGAGACAGCTGGGATG 902



```

Oy 985 TTGGCAGAGGTCCTGATGATGAGCGGTCGACCTTTATATCATGGAGACACTCTCAC 1044
Db 903 TTGGCAGAGGTCCTGATGATGAGCGGTCGACCTTTATATCATGGAGACACTCTCAC 962
Oy 1045 TTATTTAAAGATCTTGGGCTGCGTCCAAAGAGACTCTGCGGCTGGTCTGAGCTGAG 1104
Db 963 TTATTTAAAGATCTTGGGCTGCGTCCAAAGAGACTCTGCGGCTGGTCTGAGCTGAG 1022
Oy 1105 AAGAAACAAGGAGAGTGGTGGCTTCCAGTATATCATGATACAAAGTAAATATTTTCCA 1164
Db 1023 GAGAAACAAGGAGAGTGGTGGCTTCCAGTATATCATGATACAAAGTAAATATTTTCCA 1082
Oy 1165 ACTACAGTCTGGTATGAGAGTCTGACGAGAAACCTTTACCACTGGGCTGCATTTCA 1224
Db 1083 ACTACAGTCTGGTATGAGAGTCTGACGAGAAACCTTTACCACTGGGCTGCATTTCA 1142
Oy 1225 CTGGCAGTGAAGAAAGCCAGGCGCATCATGAGAGAGTTATGAGCTGTGACGCCCTCA 1284
Db 1143 CTGGCAGTGAAGAAAGCCAGGCGCATCATGAGAGAGTTATGAGCTGTGACGCCCTCA 1202
Oy 1285 ATATCACTCAGGTCTTGAAGCCATGAGAGAGGACACATCACTTTTGGATCCAAAGCTG 1344
Db 1203 ATATCACTCAGGTCTTGAAGCCATGAGAGAGGACACATCACTTTTGGATCCAAAGCTG 1262
Oy 1345 GAGTCCCTGAGACCACTCTACTTGAATGACTTATACAAAGTATTTCTTCTCCATCACTCC 1404
Db 1263 GAGTCCCTGAGACCACTCTACTTGAATGACTTATACAAAGTATTTCTTCTCCATCACTCC 1322
Oy 1405 ACGGAGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1464
Db 1323 ACGGAGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1382
Oy 1465 CTGTTGTTTCTTATGTTGTCAGACATGAGAGAAATGCTGCTGATGCTTGAAGAAAGCT 1524
Db 1383 CTGTTGTTTCTTATGTTGTCAGACATGAGAGAAATGCTGCTGATGCTTGAAGAAAGCT 1442
Oy 1525 AAGAAAGAAGAGCTTTTCACTGCTTGGGACAGAACTCTGGGCTGCAACTTTGAGAAAGCT 1584
Db 1443 AAGAAAGAAGAGCTTTTCACTGCTTGGGACAGAACTCTGGGCTGCAACTTTGAGAAAGCT 1502
Oy 1585 CCTCTTCACTAATCAATTTTCACTCAATTTTCACTCAATTTTCACTCAATTTTCACTCTT 1644
Db 1503 CCTCTTCACTAATCAATTTTCACTCAATTTTCACTCAATTTTCACTCAATTTTCACTCTT 1562
Oy 1645 CTGTTATATATCTTTCTTGAATCTTTTCAATTTCTCTGATTTCTGAGAGAGAGAGATCTCT 1704
Db 1563 CTGTTATATATCTTTCTTGAATCTTTTCAATTTCTCTGATTTCTGAGAGAGAGATCTCT 1622
Oy 1705 CCCCTCCCTCCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1764
Db 1623 CCCCTCCCTCCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1682
Oy 1765 TATATCACTCTTTAAACATTTTTCACCTTTAAAGTAAACCTTTAAATATTTTGG 1824
Db 1683 TATATCACTCTTTAAACATTTTTCACCTTTAAAGTAAACCTTTAAATATTTTGG 1742
Oy 1825 AAGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1881
Db 1743 AAGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1769

```

```

AUTHORS Liu, C. H., Lin, B. Y., and Chang, L. Y.
TITLE Cloning of the human aminopeptidase gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1472)
AUTHORS Liu, C. H., Lin, B. Y., and Chang, L. Y.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia
Sinica, RM 433, 128, Yen-Chiun-Yuan Road SEC 2, Taipei 11529, Taiwan

FEATURES
Source
1..1472
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="LCH1"
/tissue_type="liver"
/dev_stage="fetus"
54..1472
/feature="secretory protein"
/codon_start=1
/product="aminopeptidase"
/protein_id="P40321.1"
/db_xref="GI:5442030"
/translation="MKFLIFAFEGVHLSLCSKAICKNGISKRTREBIKEIASCG
DVAKAIINLAVYGAQNRYSERLALVDIVGPRLSGKNLEKAIQWYQWQDGLK
VHLEPVIPIHMERBESBAMLEPRIRHIALILGLSSISGTPEBGTAVLVVTSFDELQ
RRABEARGKIVVYNQPIYNSRTVGTGAVEAKAGALASLRSVAFSISYSPHTG
IOEVODVPKIPACITVEDABEMSRBASGKIKIVLQKGAATYPTDTSFNTVAEIT
GSKRPEQVVLVSGHSDSWDQAMMDGGAFLISWEALSLIKDGLAPKRTLRVLT
AEEGGVGAFOYIOLHKVNISSNYSLVNEDNGTFLPGLQPTSGEKRAIMEBMSLL
QPLNITGVLSHGEQDINFWIQAQVPAASLDDLKYFFPHSHGDTMTWDPKQNV
AAAVWAVSVVVDMEEMLPSS"

BASE COUNT 411 a 325 c 381 g 355 t
ORIGIN
Query Match 79.2%; Score 1465.6; DB 9; Length 1472;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1468; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Oy 46 GTGAGAGCGGCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 105
Db 1 GTGAGAGCGGCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
Oy 106 TCTTATCTTGGCAATTTTCCGTTGTTTCACTTTTATCCCTGTGCTCTGGGAAAGCTA 165
Db 61 TCTTATCTTGGCAATTTTCCGTTGTTTCACTTTTATCCCTGTGCTCTGGGAAAGCTA 120
Oy 166 TATGCAAGATGCAATCTTATGAGAGACTTTTGAAGAAATTAAGAGAAATAGCCAGCT 225
Db 121 TATGCAAGATGCAATCTTATGAGAGACTTTTGAAGAAATTAAGAGAAATAGCCAGCT 180
Oy 226 GTGAGAGTGTGTAAGGATCATCAAGCTAGCTGTTATGTAAGCCAGAGACAGAT 285
Db 181 GTGAGAGTGTGTAAGGATCATCAAGCTAGCTGTTATGTAAGCCAGAGACAGAT 240
Oy 286 CCTATGAGCGATGAGCACTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 345
Db 241 CCTATGAGCGATGAGCACTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 300
Oy 346 ACTTGAAGAAAGCCATCAATTTATGTCAGAAACCTGACAGCAAGATGAGCTGAGAAAG 405
Db 301 ACTTGAAGAAAGCCATCAATTTATGTCAGAAACCTGACAGCAAGATGAGCTGAGAAAG 360
Oy 406 TTCACTGGAGCGAGTGAATATCCCACTGGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
Db 361 TTCACTGGAGCGAGTGAATATCCCACTGGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Oy 466 TGAAGCAAGATTCATAGATAGCACTCTGGGCTTGGAGCAGCAGATTGGAGCTCTC 525
Db 421 TGAAGCAAGATTCATAGATAGCACTCTGGGCTTGGAGCAGCAGATTGGAGCTCTC 480
Oy 526 CAGAAAGCATTAACAGCAGAGAGTCTGTTGATGATGATGATGATGATGATGATGATGATGATG 585
Db 481 CAGAAAGCATTAACAGCAGAGAGTCTGTTGATGATGATGATGATGATGATGATGATGATGATG 540

```



QY 586 CTTCAGAGCAAGAGGAGATTGTTTAAACCACTTACATCACTACTCAAGCA 645  
 DB 541 CCTCAGAGCAAGAGGAGATTGTTTAAACCACTTACATCACTACTCAAGCA 600  
 QY 646 CCGTGCAATACCCGAGGAGGCGGTGGAAGCTGCCAGAGTGGGGCTTTGGCATTC 705  
 DB 601 CCGTGCAATACCCGAGGAGGCGGTGGAAGCTGCCAGAGTGGGGCTTTGGCATTC 660  
 QY 706 TCATTGATCCGAGGCTCTCTTCATCTACAGTCCCAAGATTCAGGAATACC 765  
 DB 661 TCATTGATCCGAGGCTCTCTTCATCTACAGTCCCAAGATTCAGGAATACC 720  
 QY 766 AGATGCGTGCAGCAAAATTCACAGCTGTATTAAGGTGGAAGATGCAAAATGAT 825  
 DB 721 AGATGCGTGCAGCAAAATTCACAGCTGTATTAAGGTGGAAGATGCAAAATGAT 780  
 QY 826 CAGAAATGCGTCTCTCATGAGGATCAAAATGTCATTACGCTAAGATGGGGCAAGACT 885  
 DB 781 CAGAAATGCGTCTCTCATGAGGATCAAAATGTCATTACGCTAAGATGGGGCAAGACT 840  
 QY 886 ACCGAGATCTGATCTCTCAACCTGTAGAGATCACTGGAGCAAAATTCAGAAC 945  
 DB 841 ACCGAGATCTGATCTCTCAACCTGTAGAGATCACTGGAGCAAAATTCAGAAC 900  
 QY 946 AGGTTGACTGTCAGTGAAGATCTGACAGCTGGAGATGTTGGCAGGGTCCATGATG 1005  
 DB 901 AGGTTGACTGTCAGTGAAGATCTGACAGCTGGAGATGTTGGCAGGGTCCATGATG 960  
 QY 1006 ATGGCGGTGAGCCTTTATATCATGGGAAGCATCTTATTAAGATTTGGGCTGC 1065  
 DB 961 ATGGCGGTGAGCCTTTATATCATGGGAAGCATCTTATTAAGATTTGGGCTGC 1020  
 QY 1066 GTCCAAAGAGGACTCTGGGCTGTGCTGGAAGTGAAGCAAGGTGGAGTTGGT 1125  
 DB 1021 GTCCAAAGAGGACTCTGGGCTGTGCTGGAAGTGAAGCAAGGTGGAGTTGGT 1080  
 QY 1126 CTTTCCAGTATTTATCACTACACAGGTAATATTTTCAACTACAGTCTGGTGAAGT 1185  
 DB 1081 CTTTCCAGTATTTATCACTACACAGGTAATATTTTCAACTACAGTCTGGTGAAGT 1140  
 QY 1186 CTGAGCAGAGAACTTTTCAACCTGCTGCAATTCATCTGGCAGTGAAGAGCCAGG 1245  
 DB 1141 CTGAGCAGAGAACTTTTCAACCTGCTGCAATTCATCTGGCAGTGAAGAGCCAGG 1200  
 QY 1246 CCATCATGAGAGGTTATGAGCTGTGCAAGCCCTCAATATCATCTAGTCTTGAGCC 1305  
 DB 1201 CCATCATGAGAGGTTATGAGCTGTGCAAGCCCTCAATATCATCTAGTCTTGAGCC 1260  
 QY 1306 ATGAGAGAGGAGACAGATCACTTTTGAATCCAAAGTGAAGTCTGGAGCCAGTCTAC 1365  
 DB 1261 ATGAGAGAGGAGACAGATCACTTTTGAATCCAAAGTGAAGTCTGGAGCCAGTCTAC 1320  
 QY 1366 TTGATGACTTATACAGATTTCTTCTTCATCACTCCCAAGAGACCATGATCTGCA 1425  
 DB 1321 TTGATGACTTATACAGATTTCTTCTTCATCACTCCCAAGAGACCATGATCTGCA 1380  
 QY 1426 TGAATCCAAAGCAGATGAATGTTGCTGCTGTTTGGGCTGTTTCTTATGTTG 1485  
 DB 1381 TGAATCCAAAGCAGATGAATGTTGCTGCTGTTTGGGCTGTTTCTTATGTTG 1440  
 QY 1486 CAGACATGAGAAATGCTGCTAGTCTGCTAG 1517  
 DB 1441 CAGACATGAGAAATGCTGCTAGTCTGCTAG 1472

RESULT 9  
 BC037067 1796 bp mRNA linear ROD 16-APR-2003  
 LOCUS Mus musculus plasma glutamate carboxypeptidase, mRNA (cDNA clone  
 DEFINITION MGC:46857 IMAGE:4989536), complete cds.  
 ACCESSION BC037067  
 VERSION BC037067.1 GI:22477497

## KEYWORDS

MGC.  
 Mus musculus (house mouse)

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euteleia; Rodentia; Sciurognathia; Muridae; Murinae; Mus.  
 1 (bases 1 to 1796)

## REFERENCE

1 (bases 1 to 1796)  
 Krausberg, R.L., Feltingold, E.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J.M., Hsieh, F.,  
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Schaefer, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S.,  
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Muliyil, S., Bosak, S.A., McMan, P.J.,  
 McKernan, K.J., Malek, J.A., Gamaral, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, T., Myers, R.M.,  
 Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,  
 Schermer, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 1796)  
 Krausberg, R.  
 Direct Submission  
 Submitted (23-NOV-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

## TITLE

Submitted (23-NOV-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

## JOURNAL

22388257  
 12477932  
 2 (bases 1 to 1796)

## REFERENCE

Submitted (23-NOV-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

## AUTHORS

Submitted (23-NOV-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

## JOURNAL

Submitted (23-NOV-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

## REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: nisc\_mgc@hgti.nih.gov  
 Email: gcgabs-remail.nih.gov

## COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: nisc\_mgc@hgti.nih.gov  
 Email: gcgabs-remail.nih.gov

## FEATURES

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>  
 Series: IRAP Plate: 81 Row: 9 Column: 18  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 9055233.  
 Location/Qualifiers

1. 1796  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="MGC:46857 IMAGE:4989536"  
 /tissue\_type="Colon, normal, 5 month old male mouse."  
 /clone\_id="NCI\_GAP\_C024"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 1. 1796  
 gene

/gene="Pgcp-pending"  
 /note="synonyms: H1S2, 1190003p12R1k"  
 /db\_xref="locusid:54381"  
 /db\_xref="MGI:1889205"  
 63..1475  
 /codon\_start=1  
 /product="Plasma glutamate carboxypeptidase"  
 /protein\_id="AAH37067.1"  
 /db\_xref="GI:22477498"  
 /db\_xref="locusid:54381"  
 /translation="MRSLLFLPIVHLALSGKAVFKNGVQRTFREIKERIANYEDV  
 AKALINLAVGKYQNSRYERLGLVDIVGPPSLSGSKLEKAIQIMYOLQODGLEHVA  
 LEQVRIPIHMERGESBVALEBRPHMALILGSSIGTEPPGSIIVLVAISFDEIQR  
 ASEARGLIYNQPYTEKYETVOYVQVAAKAVASLIQVASFISVSPHTGIO  
 KYODGPKRIPIACTIVEDAMMSRMSKGNITVIHLEMGARTYDQTSFNVARTGS  
 MYPERVIVUSGHDSMDVYQGLDDGGAFISMELSLVKRLGRPKRTLLVMTAE  
 EQGIGASQYVELKANTISKYSLVNEADSGTFLLPGILOFTGSDARALMKVNNLQ  
 LNVTKVFNSEGTIDINFWIQGVPSLSRLDLYKFFPHSHSGDTMTVMDKQNVVA  
 AVMAVVAIVVDMDEMLPRS"

BASE COUNT 526 a 372 c 418 g 480 t  
 ORIGIN

Query Match: 62.9%; Score 1164.8; DB 10; Length 1796;  
 Best Local Similarity 83.8%; Pred. No. 1.6e-293;  
 Matches 1329; Conservative 0; Mismatches 254; Indels 3; Gaps 1;

131 TGTTCACCTTTTATCCCTGCTGCTGGAAGCTATATGCAAGATGGCATCTTAAAG 190  
 89 TGTTCACCTTTTATCCCTGCTGCTGGAAGCTATATGCAAGATGGCATCTTAAAG 148  
 191 GACTTTGAGAAATTAAGAAATAGCCAGCTGTGAGATGTTGCTAAAGCATAT 250  
 149 AACATTTGAGAAATTAAGAAATAGCCAGCTGTGAGATGTTGCTAAAGCATAT 208  
 251 CAACCTAGCTTTTATGTTAAAGCCAGACAGATCTTATGAGCGATTGCTTCTGT 310  
 209 CAACCTAGCTTTTATGTTAAATACAGAACCGGCTCTATGAGCGTTTGGGACTTCTAGT 268  
 311 TGATACCTTTGAGAACCTGAGTGTGCTCAAGAACTTAGAAAGCATCCAAATAT 370  
 269 TGATACCTTTGAGAACCTGAGTGTGCTCTAAGAACTTAGAAAGCATCCAAATAT 328  
 371 GTACCAAAACCTGACAGATGAGGCTGAGAAAGTTCACTGAGACGATGAGATAT 430  
 325 GTACCAAAACCTGACAGATGAGGCTGAGAAAGTTCACTGAGACGATGAGATAT 388  
 431 CCACCTGAGAGAGGAGAAATCAGCTGTGATGCTGAGCCCAAGATTCATTAAGATAGC 490  
 385 CCACCTGAGAGAGGAGAAATCAGCTGTGATGCTGAGCCCAAGATTCATTAAGATAGC 448  
 491 CATCTGCTGCTTGGACACAGATGAGGATCTCCCAAGAGCATTAAGAGAAATGCT 550  
 445 TATCTTAGTCTTGGACACAGATGAGGATCTCCCAAGAGCATTAAGAGAAATGCT 508  
 551 GGTGAGTACCTCTTTCGATGAACTGACAGAAAGGCTCAGAAAGCAAGAGAAATGCT 610  
 505 GGTGAGTACCTCTTTCGATGAACTTCAAGAAAGGATCAGAAAGCAAGAGAAATGCT 568  
 611 TGTTTAATACCAACCTTACATCACTACTCAAGACGAGTGAATACCAAGAGAGGAGC 670  
 569 TGTTTAATACCAACCTTACATCACTACTCAAGACGAGTGAATACCAAGAGAGGAGC 628  
 671 GGTGAGTACCTCTTTCGATGAACTGACAGAAAGGCTCAGAAAGCAAGAGAAATGCT 730  
 629 TGTGAGTACCTCTTTCGATGAACTGACAGAAAGGCTCAGAAAGCAAGAGAAATGCT 688  
 731 CATCTAGTACCTCTTTCGATGAACTTCAAGAAAGGATCAGAAAGCAAGAGAAATGCT 790  
 689 CATCTAGTACCTCTTTCGATGAACTTCAAGAAAGGATCAGAAAGCAAGAGAAATGCT 748  
 791 AGCTGTATTAACGTTGAGAAATGCAAGAAATGATGCAAGAAATGCTTCTCAATGAGATCA 850  
 749 AGCTGTATTAACGTTGAGAAATGCAAGAAATGATGCTGAAATGCTTCTCTGGGAGCAAA 808

851 AATTGCTATTCAGCTAAAGATGGGGGCAAGAACTTACCAAGATATGATTCCTTCAAC 910  
 809 AATTGCTATTCATTTGGAAATGGGAGCAAGAACTTACCAAGATATGATTCCTTCAAC 868  
 911 TGTAGCAGAGATCACTGGAGCAAAATATTCAGAAACAGGTTGTATGCTGATGAGATCT 970  
 869 TGTTCAGAGATCACTGGAGCAAGATATTCAGAGAAAGTGTCTGCTGATGAGATCT 928  
 971 GGAAGCTGGAGATTTGGAGAGGCTGCAATGATGATGAGGCTGATGATGATGATGATG 1030  
 929 GGAAGCTGGAGATTTGGAGAGGCTGCAATGATGATGAGGCTGATGATGATGATGATG 988  
 1031 GGAAGCTGGAGATTTGGAGAGGCTGCAATGATGATGAGGCTGATGATGATGATGATG 1090  
 989 GGAAGCTGGAGATTTGGAGAGGCTGCAATGATGATGAGGCTGATGATGATGATGATG 1048  
 1091 GCTCTGAGCTGCAAGAAACAGAGTGGAGTGTGCTTCCAGTATTTATGATTTACAA 1150  
 1049 GCTCTGAGCTGCAAGAAACAGAGTGGAGTGTGCTTCCAGTATTTATGATTTATGAT 1108  
 1151 GGTAAATATTTTCCATCACTGCTGCTGATGAGTCTGACGCAAGAACTTTTATCCAC 1210  
 1109 GGTAAATATTTTCCATCACTGCTGCTGATGAGTCTGACGCAAGAACTTTTATCCAC 1168  
 1211 TGGGCTGCAATTCATGCGAGTGAAGGCAAGGCGCCATCATGAGAGGATTTATGAGCCT 1270  
 1169 TGAATCTGATTCATGCGAGTGAAGGCAAGGCGCCATCATGAGAGGATTTATGAGCCT 1228  
 1271 GCTGACAGGCTTCAATTCATGCTGCTGATGAGTCTGACGCAAGAACTTTTATCCAC 1330  
 1229 TCTGCAAGGCTTCAATTCATGCTGCTGATGAGTCTGACGCAAGAACTTTTATCCAC 1288  
 1331 TTTGATTCAGAGTGTGAGTGTGCTGAGGCAAGTCTTATGATGATGATGATGATGATG 1390  
 1289 CTGATTCAGAGTGTGAGTGTGCTGAGGCAAGTCTTATGATGATGATGATGATGATG 1348  
 1391 CTGATTCAGAGTGTGAGTGTGCTGAGGCAAGTCTTATGATGATGATGATGATGATG 1450  
 1349 TTTGATTCAGAGTGTGAGTGTGCTGAGGCAAGTCTTATGATGATGATGATGATGATG 1408  
 1451 TGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1510  
 1409 TGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1468  
 1511 GTCTTGAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1570  
 1469 GTCTTGAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1528  
 1571 AACTTGAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1630  
 1529 AACTTGAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1588  
 1631 AACTTGAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1690  
 1589 TTTCTTATACCTTCTGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1645  
 1691 AAGGAATCATTTCTCCCTCCCTCC 1716  
 1646 TAAATCATGATTCCTACTGACACACC 1671

RESULT 10  
 AX400829  
 LOCUS AX400829 1778 bp DNA linear PAT 06-JUN-2002  
 DEFINITION Sequence 505 from Patent WO0210453.  
 ACCESSION AX400829  
 VERSION AX400829.1 GI:21337009  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1  
AUTHORS Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and  
TITLE Elashoff, M.R.  
JOURNAL Molecular toxicology modeling.  
Gene Patent: WO 0210453-A 505 07-FEB-2002;  
FEATURES  
source location/Qualifiers  
1. 1778  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/note="EMBL/GenBank Accession No. AF097723"  
BASE COUNT 488 a 391 c 442 g 457 t  
ORIGIN  
Query Match 61.9%; Score 1145; DB 6; Length 1778;  
Best Local Similarity 81.9%; Pred. No. 2.4e-288;  
Matches 1334; Conservative 0; Mismatches 290; Indels 5; Gaps 1;  
QY 73 AACAGAAACCACTGAGAAAAAATGAAATTCCTATCTTGCCATTTTCCGTGGT 132  
Db AGCAAGAGAAAGAACTAGGACATAGAGTTCCTTCTCTGTCGTGCTGTTG 154  
QY 133 TTCACTTTTATCCCTGCTCTGCGAAGCTATATGCAAGATGCGATCTTAAGAGA 192  
Db 155 TTCACTTTTCTCTGCGGCTCTGGAAGAGCTATATACAGAGTGTCTTCAAGCAA 214  
QY 193 CTTTGAAGAAATTAAGAGAAATAGCCAGCTGTGAGATGTGCTAAGCATCA 252  
Db 215 CATTTCAAGAAATTAAGAGAAATAGCCAACTAAGAGATGTGCTAAGCAATTATCA 274  
QY 253 ACCAGCTGTTATGCTTAAAGCCAGAACAGATCTATAGAGATGCGACTTCGTGG 312  
Db 275 ACCTGCTGTTATGAGAAATACAGAACCGTCTATAGAGCTTTGAGACTTCTAGTTG 334  
QY 313 ATACTGTGAGACCGAGCTGAGTGGCTCAAGAACTTGAAGAAAGCCATCCAAATTATGT 372  
Db 335 ATACTGTGAGACCGAGCTGAGTGGCTTAAAGAACTTGAAGAAAGCTTCCAAATCATGT 394  
QY 373 ACCAAACCTGACAGAGATGGGCTGAGAAAGTTACCTGAGGCACTGAGAAATACCC 432  
Db 395 ACCAAACCTGACAGAGATGGGCTGAGAAAGTTACCTGAGGCACTGAGAAATACCTC 454  
QY 433 ACTGGAGAGGGGAGAAATACAGTGTGATGCTGAGACCCAGAAATTCATAAGATAGCA 492  
Db 455 ACTGGAGAGGGGAGAAATTCGAGTGTGATGCTGAGTGTGCTGAGTGTGCTGAGTGTG 514  
QY 493 TCTGGGCTCTGGAGCAGCATTTGGGACTCTCCAGAAAGCATTCACAGAGATTTCTGG 552  
Db 515 TTTTAGGCTTGGCGGAGCATTTGGGACTCTCTCGAAGGATTCACAGAGATTTCTGG 574  
QY 553 TGTGACCTCTTTCGATGATGCTGAGAGAGGGCTCGAAGCAAGAGGAAAGATTGTG 612  
Db 575 TGTGACCTCTTTCGATGATGCTGAGAGAGGGCTCGAAGCAAGAGGAAAGATTGTG 634  
QY 613 TTTTATACCAACTTACATCACTACTCAAGAGCGGTGCAATCCGAAAGCGAGGGGGGG 672  
Db 635 TTTTATACCAAGCTTACCTGATGAGAAACGTGCAAGTACCGGAGCGCGAGGCTG 694  
QY 673 TGAAGCTGCGCAAGGTGGGGGCTTGGCATCTCTGATTCGATCCGTGGCTCTCTCTCA 732  
Db 695 TGAAGCTGCGCAAGGTGGGGGCTTGGCATCTCTGATTCGATCCGTGGCTCTCTCTCA 754  
QY 733 TCTACAGTCTCTACACAGGTATTCAGAGATACAGAGATGCGTCCAAATTCACAGAG 792  
Db 755 TCTACAGTCTCTACACAGGTATTCAGAGATACAGAGATGCGTCCAAATTCACAGAG 814  
QY 793 CCTGATTAAGGTGAGAGATGAGAAATGATGTCAAGATGCTTCTCATGGGATCAAAA 852  
Db 815 CCTGATTAAGGTGAGAGATGAGAAATGATGTCTCATGGGATCAAAA 874  
QY 853 TTGATATTGAGTAAAGATGGGGGCAAGACCTACCCAGATTAAGTATTCCTTCAACACTG 912

Db 875 TTGATATTGAGTAAAGATGGGGCAAGACCTTACAGATTAAGATTCCTTCAACACTG 934  
QY 913 TAGCAGAGATCATCTGGAGAGCAAAATTCAGAAACAGTTGATCTGATGAGATCTGG 972  
Db 935 TTGAGAGATCATCTGGAGAGCAAAATTCAGAAACAGTTGATCTGATGAGATCTGG 994  
QY 973 ACAGCTGGAGATGTTGGGAGGAGTCCAGATGATGATGAGGAGGAGCTTTATCATGGG 1032  
Db 995 ACAGCTGGAGATGTTGGGAGGAGTCCAGATGATGATGAGGAGGAGCTTTATCATGGG 1054  
QY 1033 AAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGACTCTGCGGTGTGC 1092  
Db 1055 AAGCACTCTCACTTATTAAGATCTTGGGCTGCGTCCAAAGAGACTCTGCGGTGTGC 1114  
QY 1093 TCTGAGCTGCAAGAGAAACAGGTGAGTGTGCTCCAGATTTATCATCAAGAG 1152  
Db 1115 TCTGAGCTGCAAGAGAAACAGGTGAGTGTGCTCCAGATTTATGAGCTTATCAAGG 1174  
QY 1153 TAAATATTTCCAACTACAGTCTGAGTGTGAGTGTGAGAGCAAGCACTTACCACTG 1212  
Db 1175 CAATATTTCCAACTACAGTGTGAGTGTGAGTGTGAGAGCAAGCACTTACCACTG 1234  
QY 1213 GGTGCAATTCATCTGCAAGTGAAGAAAGCCAGGCAATCATGAGAGGTTATGAGCCTGC 1272  
Db 1235 GGTGCAATTCATCTGCAAGTGAAGAAAGCCAGGCAATCATGAGAGGTTATGAGCCTGC 1294  
QY 1273 TGCAGCCCTCTAATATCATCTAGTCTGAGCCATGAGAAAGGACAGACATCACTTT 1332  
Db 1295 TGCAGCCCTCTAATATCATCTAGTCTGAGCCATGAGAAAGGACAGACATCACTTT 1354  
QY 1333 GGATCCAACTGAGAGTGTGAGGAGCCAGTCTACTGATGATGATTAACAAGTATTTCTCT 1392  
Db 1355 GGATCCAACTGAGAGTGTGAGGAGCCAGTCTACTGATGATGATTAACAAGTATTTCTCT 1414  
QY 1393 TCCATCACTCCCAAGAGACACATGATCTGATGATGATGATGATGATGATGATGATG 1452  
Db 1415 TCCATCACTCCCAAGAGACACATGATCTGATGATGATGATGATGATGATGATGATG 1474  
QY 1453 CTGCTGTTGGGCTGTTGTTCTTATGTTGTTGAGACATGAGAAATGCTGCTAGGT 1512  
Db 1475 CTGCTGTTGGGCTGTTGTTCTTATGTTGTTGAGACATGAGAAATGCTGCTAGGT 1534  
QY 1513 CCTAGAAACAGTAAAGAAAGAAAGCTTTTCACTTCTGAGGAGAAATCTGGGCTGCA 1572  
Db 1535 CCTAGAAACAGTAAAGAAAGAAAGCTTTTCTGAGGAGAAATCTGGGCTGCA 1594  
QY 1573 CTTGGAAGAACTCTCTTCACTAATCAATTCATCAATTCATCAATTCATCAATTCAT 1632  
Db 1595 TTTCAAGCAGCAGCATCTTCAAGCAGCAGTGTATTAACATCAATCCCGGTGACAC 1649  
QY 1633 ATTTGATGCTTCTGTTATTAATCTTTCTTGAATCTTTCAAAATCTGATTTAGAAA 1692  
Db 1650 GTTCTTATTAATCTTCTGTTATTAATCTTTCTTGAATCTTTCAAAATCTGATTT 1709  
QY 1693 AGGAATCAT 1701  
Db 1710 AGTAATCAT 1718  
RESULT 11  
AF097723  
LOCUS 1778 bp mRNA linear ROD 10-NOV-1998  
DEFINITION Rattus norvegicus hematopoietic lineage switch 2 related protein  
(H1S2-1P) mRNA, complete cds.  
ACCESSION AF097723  
VERSION AF097723.1 GI:3851631  
KEYWORDS  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 1778)  
AUTHORS Chen, Y., and Talmage, D.  
TITLE Direct Submission  
JOURNAL Submitted (07-OCT-1998) Institute of Human Nutrition, Columbia  
University, 701 West 168th Street Room 5-503, New York, NY 10032,  
USA

FEATURES  
source Location/Qualifiers

1..1778  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Fisher"  
/db\_xref="taxon:10116"  
/cell\_line="F11"  
/cell\_type="fibroblast"  
/dev\_stage="embryo"  
1..1778  
/gene="H1a2-rp"  
121..1539  
/gene="H1a2-rp"  
/codon\_start=1  
/product="hematopoietic lineage switch 2 related protein"  
/protein\_id="AAC72384.1"  
/db\_xref="gi:3851632"  
/translation="MRFLEFLVAVVHLFSLGSKAIYKSVQRTQEIKEIANYE  
DVAKIINLAVYKXONRYERLGLVDVGPRLSGSKLEKAIQIMYONQDGLN  
VHLEQVRIPHWGSESAVMVPRILKALIGLGGISITPEEGITAEVLVVASFEIQ  
RASEARGKIVVNOPTDYKTVQYRGRVAEAKVAVASLRSVASFISVPHG  
HOGYODGVPIPTACITTEDAEMMSRMSRBDKIVILKMGAKYPTDSSNTYAEIT  
GSKYPERVVLVSHLDSMDVOCALDDCGAFISWEALSLVKDLGRKRLVLVLT  
AEQGVGASQYELHKNANISKYSLVMSADSGTLPGLQTSDDKAKAIKEVWSLW  
QPNATIKFNDABGTDINFMIOAGVPGASLSDLDLYKFFHSHSDTMTADPKOMY  
AAAVWAVVAVVADMEMLPRS"

BASE COUNT 488 a 391 c 442 g 457 t  
ORIGIN

Query Match 61.9%; Score 1145; DB 10; Length 1778;  
Best Local Similarity 81.9%; Pred. No. 2,4e-288;  
Matches 1334; Conservative 0; Mismatches 290; Indels 5; Gaps 1;

```
73 AACAGAAACCACTGGAAGAAAAAATGAATTCCTATCTTCGATTTTCGTGTG 132
Db AGCAGAGAAAGAAAGAACTAGGACATGAGGTCCTTTCTCTGTTCGTGTG 154
95 AGCAGAGAAAGAAAGAACTAGGACATGAGGTCCTTTCTCTGTTCGTGTG 154
133 TTCACTTTTATCCCTGTCTCTGGGAAGCTATATGCAAGATGGCATCTTAAGAGA 192
Db TTCACTTTTCTCTGTGGCTCTGGAAGAGCTATATCAAGAGTGTCTTCTACGCA 214
155 TTCACTTTTCTCTGTGGCTCTGGAAGAGCTATATCAAGAGTGTCTTCTACGCA 214
193 CTTTGAAGAAATGAAGAAATAGCCAGCTGTGAGATGTTGCTAAGCAATCATCA 252
Db CATTTCAGAAATGAAGAAATAGCCAACTATGAGATGTTGCTAAGCAATCATCA 274
215 CATTTCAGAAATGAAGAAATAGCCAACTATGAGATGTTGCTAAGCAATCATCA 274
253 ACCTAGCTTTTATGGTAAAGCCAGACAGATCCTATGAGAGTTGGCATCTCGTGTG 312
Db ACCTGCTGTTTATGGAATAACAGAACCGGCTCTATAGAGTTGGACCTTCAAGTG 334
275 ACCTGCTGTTTATGGAATAACAGAACCGGCTCTATAGAGTTGGACCTTCAAGTG 334
313 ATACTGTTGACCCAGACTGAGTGGCTCAAGAACTTAGAAAAAGCCATCCAAATTATGT 372
Db ATACTGTTGACCCAGACTGAGTGGCTCTAGAACCTTAGAAAAAGCTATCCAAATCATGT 394
335 ATACTGTTGACCCAGACTGAGTGGCTCTAGAACCTTAGAAAAAGCTATCCAAATCATGT 394
373 ACCAAACCTGACAGAGATGGCTGGAAGAAATTACCTGAGGCACTGAGAAATACCC 432
Db ACCAAACCTGACAGAGATGGCTGGAAGAAATTACCTGAGGCACTGAGAAATACCTC 454
395 ACCAAACCTGACAGAGATGGCTGGAAGAAATTACCTGAGGCACTGAGAAATACCTC 454
433 ACTGGAGAGGGGGAAGAAATACAGCTGATGATGCTGAGACCAAGATTGTAAGATAGCA 492
Db ACTGGAGAGGGGGAAGAAATCTGCAATGATGATGCTGAGACCAAGATTGTAAGATAGCA 514
455 ACTGGAGAGGGGGAAGAAATCTGCAATGATGATGCTGAGACCAAGATTGTAAGATAGCA 514
493 TCTGGGCTCTTGGAGCAGATTTGGGACTCTCCAGAGGCAATTACAGACAGAGTTCTGG 552
Db TTTTAGGCTTGGGAGCAGATTTGGGACTCTCTGAGAGGTAATCAACAGAGATCTCG 574
515 TTTTAGGCTTGGGAGCAGATTTGGGACTCTCTGAGAGGTAATCAACAGAGATCTCG 574
553 TGTGACCTCTTTCGATGAATGACAGAGAGGCTCTGAGAGCAAGAGAGAAATGTTTG 612
|||||
```

```
575 TGTGAGCTCTTTTGTGAACCTTCAAGAGGCGATCAGAGCAAGAGGAATGTTG 634
613 TTTATAACCAACCTTATCATCACTACTCAAGAGAGGTCATACGAGACGAGGAGCGG 672
Db TTTATAACCAACCTTATCATCACTACTCAAGAGAGGTCATACGAGAGGAGCGGAGCTG 694
635 TTTATAACCAACCTTATCATCACTACTCAAGAGAGGTCATACGAGAGGAGCGGAGCTG 694
673 TGAAGCTGCCAAGTGGGGCTTTGGCATCTCTCATTCGATCCGTCCTTCTCA 732
Db TGAAGCTGCCAAGTGGGGCGGTGGCATCTCTCATTCGATCCGTCGATGCTTTTCA 754
695 TGAAGCTGCCAAGTGGGGCGGTGGCATCTCTCATTCGATCCGTCGATGCTTTTCA 754
733 TCTACAGTCTCAGACAGGATTTAGAAATACAGAGATGGCGTCCCAAAATTTCAACAG 792
Db TCTACAGTCTCAGACAGGATTCAGAAATATCAAGATGGGTGCCCAAGATTTCAACAG 814
755 TCTACAGTCTCAGACAGGATTCAGAAATATCAAGATGGGTGCCCAAGATTTCAACAG 814
793 CCTGTATTAACGATGGAAGATGACAGAAATGATGTCAGAAATGGCTTCTCATGGATCAAA 852
Db CCTGTATTAACGATGGAAGATGACAGAAATGATGTCAGAAATGGCTTCTCATGGATCAAA 874
815 CCTGTATTAACGATGGAAGATGACAGAAATGATGTCAGAAATGGCTTCTCATGGATCAAA 874
853 TTGTCAATTCAGCTAAAGATGGGGCAAAAGACTTACCCAGATCTGATTCCTTCAACATG 912
Db TTGTCAATTCAGCTAAAGATGGGGCAAAAGACTTACCCAGATCTGATTCCTTCAACATG 934
875 TTGTCAATTCAGCTAAAGATGGGGCAAAAGACTTACCCAGATCTGATTCCTTCAACATG 934
913 TAGCAGAGATCACTGGAGCAATATTCAGAACAGTTGTACTGTCAGTGGACATCTGG 972
Db TTGACAGAGATCACTGGAGCAAGATATCAGAGAAAGTTGTCCTGTCAGTGGACATCTGG 994
935 TTGACAGAGATCACTGGAGCAAGATATCAGAGAAAGTTGTCCTGTCAGTGGACATCTGG 994
973 ACAGCTGGAGATGTTGGGCGAGGTCAGATGATGATGAGCGGTCGACCTTATATCATGGG 1032
995 ACAGCTGGAGATGTTGGGCGAGGTCAGATGATGATGAGCGGTCGACCTTATATCATGGG 1054
1033 AAGCACTCTCACTTATTAAGATCTTGGGTCGCTCCAAAGAGACTCTGGCGCTGTGTC 1092
Db AAGCACTCTCACTTATTAAGATCTTGGGTCGCTCCAAAGAGACTCTGGCGCTGTGTC 1114
1055 AAGCACTCTCACTTATTAAGATCTTGGGTCGCTCCAAAGAGACTCTGGCGCTGTGTC 1114
1093 TCTGAGCTGCAAGAAACAAAGTGGATGTCGCTCCAGATATTCAGTATACAAAG 1152
Db TCTGAGCTGCAAGAAACAAAGTGGATGTCGCTCCAGATATTCAGTATACAAAG 1174
1115 TCTGAGCTGCAAGAAACAAAGTGGATGTCGCTCCAGATATTCAGTATACAAAG 1174
1153 TAAATATTTCAACTATACAGTCTGATGATGAGTGTGACGAGGAACCTTCTTACCACTG 1212
1175 CAATATTTTCAACTATACAGTCTGATGATGAGTGTGATGAGGCTGATCAGGAACCTTCTTACCACTG 1234
1213 GCGTCATTTCACTGCGAGTGAAGAAAGCCAGGCGCATCATGAGAGAGTTATGAGCCTGC 1272
Db GCGTCAGTTTCACTGCGAGTGAAGAAAGCCAGGCGCATCATGAGAGAGTTATGAGCCTGC 1294
1235 GCGTCAGTTTCACTGCGAGTGAAGAAAGCCAGGCGCATCATGAGAGAGTTATGAGCCTGC 1294
1273 TGCAGCCCTTCATATATATCACTCAGGCTCTGAGCCATGAGAAAGGACAGACATCACTTT 1332
Db TGCAGCCCTTCATATATATCACTCAGGCTCTGAGCCATGAGAAAGGACAGACATCACTTT 1354
1295 TGCAGCCCTTCATATATATCACTCAGGCTCTGAGCCATGAGAAAGGACAGACATCACTTT 1354
1333 GGATCCAGGCTGAGTGCCTGAGCCAGTCTACTTGAATGATCTTATCAAGATTTCTTCT 1392
Db GGATCCAGGCTGAGTGCCTGAGCCAGTCTACTTGAATGATCTTATCAAGATTTCTTCT 1414
1355 GGATCCAGGCTGAGTGCCTGAGCCAGTCTACTTGAATGATCTTATCAAGATTTCTTCT 1414
1393 TCCATCACTCCACGAGACACATGACTGTATGATTCAGATTCAGAGAGATGATTTGCTG 1452
Db TCCATCACTCCACGAGACACATGACTGTATGATTCAGATTCAGAGAGATGATTTGCTG 1474
1415 TCCATCACTCCACGAGACACATGACTGTATGATTCAGATTCAGAGAGATGATTTGCTG 1474
1453 CTGCTGTTTGGGCTGTTGTTTATATGTTGTTGACAGATGAGAAATGCTGCTAGT 1512
Db CTGCTGTTTGGGCTGTTGTTTATATGTTGTTGACAGATGAGAAATGCTGCTGAGT 1534
1475 CTGCTGTTTGGGCTGTTGTTTATATGTTGTTGACAGATGAGAAATGCTGCTGAGT 1534
1513 CCTAGAAACAGTAAAGAAAGGTTTTCANGTCTGAGCAGGAATCTCGGGCTGCA 1572
Db CCTAGAAACAGTAAAGAAAGGTTTTCANGTCTGAGCAGGAATCTCGGGCTGCA 1594
1535 CCTAGAAACAGTAAAGAAAGGTTTTCANGTCTGAGCAGGAAATCTCGGGAT 1594
1573 CTTTGAAGAACTCCTCTTCAATTAATTCATCAATTCATCTTCAAGACCACTCT 1632
Db TTTTACAGACGACATCTTCAAGACCTTGTATACACTCAATCCCGTGGACACA----- 1649
1595 TTTTACAGACGACATCTTCAAGACCTTGTATACACTCAATCCCGTGGACACA----- 1649
1633 ATTTCATGCTTTCTGTTATATCTTTTGTGATACCTTTCAAATTTCTGTATTTAGAAA 1692
Db GTTTCCTTATACCTTCTGTTATACCTTTCCTTATACCTTTCCTGTTTACGATA 1709
```

```

QY      1693 AGAATCAT 1701
      |||||
Db      1710 AGTATCAT 1718

RESULT 12
AF131077      1726 bp mRNA linear ROD 03-JAN-2003
LOCUS      Rattus norvegicus liver annexin-like protein (LAL) mRNA, complete
DEFINITION
ACCESSION  AF131077
VERSION     AF131077.1 GI:7108712
KEYWORDS   Rattus norvegicus (Norway rat)
SOURCE     Rattus norvegicus
ORGANISM   Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 1726)
            Della Fazio, M.A., Piobico, D., Bartoli, D., Castellini, M.,
            Brancorsini, S., Viola Magni, M. and Servillo, G.
            lal-1: a differentially expressed novel gene during proliferation
            in liver regeneration and in hepatoma cells
            Genes Cells 7 (11), 1183-1190 (2002)
            22278398
            12390252
REFERENCE  2 (bases 1 to 1726)
            Servillo, G., Della Fazio, M.A., Piobico, D., Bartoli, D.,
            Castellini, M., Brancorsini, S. and Viola Magni, M.
            Direct Submission
            Submitted (25-FEB-1999) Institute of General Pathology, University
            of Perugia, Policlinico Monteluce, Perugia 06100, Italy
FEATURES
    source
        1..1726
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /db_xref="taxon:10116"
            1..1726
            /gene="LAL"
            63..1481
            /note="Involved in liver regeneration; contains unique
            annexin domain at the N-terminus"
            /product="liver annexin-like protein"
            /protein_id="AAF36518.1"
            /db_xref="GI:7108713"
            /translation="MRFLEFLFVAVVHLFSLGSKAIYKSGVSRFTQEIKEETANYE
            DVAKAIINLAVYGYONRSYERGLDLDVTGPRLSGSKNEKAIQIOMYLOODPLEN
            VHLFOVRIPIHMERGESAVVWVPRILHLAILGLGSGICTPEEGITAEVILVVASFVEIQ
            RRAASBARKIVYNQPTDYTGKYQYREKAVAEAKGAVASLIRVASSTISPHFG
            HGGIDGVPIKPIACTITTEDMEMSRTASRQDKIVHLKMGAKYPTDSEFNTVAET
            GSKYPERVIVVSGHLDSDVGGALDDGGAFSLWELSLVKDGLRPKRLRLVLT
            ABERGGVGAQYVELHRANISKYSLVMEADSGTFLPTGLQFTGSDKARAIMKEVMSLL
            OPNLTIKVFADSGTIDINFWIOAGVPGASLRDLKYKFFFHSHSDMTAMDPRQMWV
            AAVMVAIVAVVADMEEMLRPS"
BASE COUNT  474 a      385 c      428 g      439 t
ORIGIN
Query Match      61.8%; Score 1143.8; DB 10; Length 1726;
Best Local Similarity 80.9%; Pred. No. 4.9e-288;
Matches 1346; Conservative 0; Mismatches 312; Indels 5; Gaps 1;

```

```

Db      123 AAAGCTATATACAAAGAGTGGTGTTCACAGGAACATTCACAGAAATTAAGAGAAATA 182
QY      219 GGCAGCTGTGAGATGTTTGTAAAGCAATCATCAACTAGTGTTTATGTTAAAGCCAG 278
Db      183 GCCAACTATGAAAGTGTGTAAAGCAATTCACACCTGTGTATTATGAAATATACAG 242
QY      279 AACGATCTATGAGCGATTGGCAGCTCTGTTGTAATCTGTGTGAGCCAGCTAGTGGC 338
Db      243 AACGGTGTGATGAGCGTTTGGAGCTTCTAGTTGATCTGTTGAGCCAGCTAGTGGC 302
QY      339 TCAGAACTTAGAAAAAGCCATCCAAATTATGTACCAAACTGACAGCAAGATGGCTG 398
Db      303 TCTAAGAACTTAGAAAGCAATTCATCAATGATGACCAAACTGCAACAAAGATGGCTG 362
QY      399 GAGAAATTCACCTGAGAGCCAGTGAAGTAATCCCATCTGGAGAGAGGAGAAATCACT 458
Db      363 GAAAACGTCAACCTGAGAGGTCAGAAATCTTCACCTGAGAGAGGAGGAGAAATCTGCA 422
QY      459 GTGATGCTGAGAGCCAAATTCATTAAGTATGACCATCTGGGCTTTGGACAGCATTTGG 518
Db      423 GTGATGCTGAGCTTCCAAATTCACAACTGCTATTTAGCCCTTGGCGGAGCATTTGG 482
QY      519 ACTCCTCAGAAAGCATTACAGCAAGAGTCTGTGTGACCTCTTTCATGATGATGACAG 578
Db      483 ACTCCTCAGAAAGTATCAGCAAGAGTCTGTGTGACCTCTTTCATGATGATGACAG 542
QY      579 AGAAGGCTCAGAGCAAGAGAGAGAGATGTTGTTTAAACCACTTACATCACTAC 638
Db      543 AGAAGGCTCAGAGCAAGAGAGAGATGTTGTTTAAACCACTTACATCACTAC 602
QY      639 TCAAGAGCGTCAATCCGAGAGCGAGCGGCGGTGAGTCTGCAAGTGGGGCTTTG 698
Db      603 GCGAAACTGTGACGACGAGAGCGGAGCGGTGAGTCTGCAAGTGGGGCTTTG 662
QY      699 GCATCTCTCATTCGATCCGTGGCTCTCTCTCATCTACAGTCTCAGACAGTATTCAG 758
Db      663 GCATCTCTCATTCGATCAGTATGAGTCTCTCTCTCATCTACAGTCTCAGACAGTATTCAG 722
QY      759 GAATACCAAGATGGCGTGGCCCAAAATTCACACAGCTGTATTAACGTGGAAGTGCAGAA 818
Db      723 GAATACCAAGATGGTGGCCCAAGATTCACACAGCTGTATTAACGTGGAAGTGCAGAA 782
QY      819 ATGATGTCAGAAATGCTTCTCATGAGTCAAAATGTCATTCAGTCAAGATGGGGCA 878
Db      783 ATGATGTCAGAAATGCTTCTCATGAGTCAAAATGTCATTCAGTCAAGATGGGGCA 842
QY      879 AAAGCTAACCAATACATGATTCCTTCAACATCTGTACAGAGATCACTGGAGAGCAATAT 938
Db      843 AAAGCTATCAAGATACAGATTCCTTCAACATCTGTACAGAGATCACTGGAGAGCAATAT 902
QY      939 CCAGAACAGTGTGATCTGTACAGTGCAGATCTGACAGCTGGAGATGTTGGCAGAGTGC 998
Db      903 CCAGAACAGATGTTGCTGTGCTGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGC 962
QY      999 ATGATATAGCGGTGAGAGCTTTATATCATGGAAGCACTTCACTTATTAAGATCTT 1058
Db      963 CTGATATAGCGGTGAGAGCTTTATATCATGGAAGCACTTCACTTATTAAGATCTT 1022
QY      1059 GGGCTGCTCAAAAGAGACTGTGGGCTGTGCTGTGACTGAGAGAGAAAGAGTGGGA 1118
Db      1023 GGGCTGCTCAAAAGAGACTGTGGGCTGTGCTGTGACTGAGAGAGAGAAAGAGTGGGA 1082
QY      1119 GTTGTGCTTCCAGATATATACATTAAGAGTAAATTTTCAACTACAGTGTGTTG 1178
Db      1083 GTTGTGCTTCCAGATATATACATTAAGAGTAAATTTTCAACTACAGTGTGTTG 1142
QY      1179 ATGAGTCTGACCAAGAACTTCTTACCACTGGGTGCAATTCATCTGCACTGGAAG 1238
Db      1143 ATGAGTCTGACCAAGAACTTCTTACCACTGGGTGCAATTCATCTGCACTGGAAG 1202
QY      1239 GCCAGGCTCATGAGAGAGTATGAGCTCTGAGAGCCCTCAATATCACTCAGGTC 1298

```

BASE COUNT	415 a	357 c	421 g	403 t
ORIGIN				
Query Match	60.4%	Score 118.4;	DB 10;	Length 1596;
Best Local Similarity	87.9%;	Pred. No. 2.2e-281;		
Matches 1218;	Conservative	0;	Mismatches 168;	Indels 0; Gaps 0
Qy	131	TGTTGACCTTTTATTCCTCTGTCCTGGGAAAGCTATATGCAAGAATGGCATCTCTTAAGG	190	
Db	210	TGTTTCACCTTTTACGCTTGGGCTCTGGTAAAGCTGTATTCAGATGCTGTTCACAG	269	
Qy	191	GACTTTTGAAGAAATTAAGAAGAAATGCGAGCTGTGAGATGTTGGCTAAAGCAATCAT	250	
Db	270	AACATTTTGAAGAAATTAAGAAGAAATGCGCAATTTATAGATGTTGGCTAAAGCAATTTAT	329	
Qy	251	CAACCTGACTGTTTATGTGTAAAGCCAGCAAGACAGATTCCTATGAGCGATTGGCACTTCTGGT	310	
Db	330	CAACCTGCTGTTTATGTGTAAATATACAGAAACCGGTCCTATGAGCGATTGGGACTTCTAGT	389	
Qy	311	TGATATCTGTTGGACCCAGACTGATGTGGTCCAAAGAACTTAAGAAAAAGCATTCCAATTTAT	370	
Db	390	TGATATCTGTTGGACCCAGACTGATGTGGTCTTAAGAACTTAAGAAAAAGCATTTCAATATCAT	449	
Qy	371	GTACCAAAACCTGCAGCAAGATGGGCTGTGAAGAAAGTTACCTGCAGCGCAGTGAAGATACC	430	
Db	450	GTACCAAAACCTGCAGCAAGATGGGCTTAAGAAATGTTTACCTGCAGCGCAGTGAAGATACC	509	
Qy	431	CCACTGGAGAGGGGAGAGAAATCAGCTGTGATGCTGGAGCCAAAGAAATTCATTAAGATAGC	490	

Qy	491	TATCTGGGCTTTGGGAGACATTTGGGACTCTCCGAGAAAGCATTAACGAGAAAGTCT	550
Db	570	TATTTCTAGGTTTGGACAGACATTTGGGACTCTCCGAGAGACATCAACGAAAGTCT	629
Qy	551	GGTGGTGAACCTCTTTCGATGAAGTGCAGAGAGAGGGCCTCAGAAGCAAGGGAAGATTGT	610
Db	630	GGTGGTAGCCTCTTTTGATGAATCTTCAAGAAGAGCATCAGAAGCAAGGGAAGATCAT	689
Qy	611	TGTTTATTAACCAACCTTATCAACTACTACAGACGGTGCATATCCGACAGCGAGGGGC	670
Db	690	TGTTTATTAACCAACCTTATCACTGGCTATGAGAAACATGTGAGTACCGGTGCAGGGAGC	749
Qy	671	GGTGAAGCTGCCAAGGTGGGGCTTTGGGACTCTCTCATTTGCATCCGTGGCCTCTTTC	730
Db	750	TGTGAAGCTGCCAAGGTGGGAGCTGTGGATCCCTCATCAATCAGTAGCTCTCTTTTC	809
Qy	731	CATTTACAGTCTCTACACAGGATATTCAGGAATACCAAGATGGCGTGCCCAAAATTTCCAC	790
Db	810	CATTTACAGTCTCTACACAGGATATTCAGGAATATCAAGATGGTGTGCCCAAGATTTCCAC	869
Qy	791	AGCTGTATTAACGGTGGAGAGTGCAGAAATATATGTCAAGATTTGGCTTCTCATGGGATCA	850
Db	870	AGCTGTATTAACGGTGGAGAGTGCAGAAATATATGTCTCGAATGGCTTCTCTGGGAAACA	929
Qy	851	AATTGTCAATTCAGTTAAAGATGGGGCAAGACCTAACCCAGATACTGATTTCTTCCACAC	910
Db	930	AATTGTCAATTCAGTTAAATGGGAGCAAAACCTATTCAGATCTGATTTCTTCCAAATAC	989
Qy	911	TGTAGCAGAGTCACTGGGAGCAAAATATCCGAACAGATTGTACTGGCAGTGGACATCT	970
Db	990	TGTTGCAGAGTCACTGGGAGCATGTATCCAGAGAAAGTTGTCTTGGTCACTGGACATTT	1045
Qy	971	GGAAGCTGGGATGTTGGGAGGGGTGCCATGGATGATGGCGGTGAGCCTTTATATCAATG	1030
Db	1050	GGAAGCTGGGATGTTGGGAGGGGTGCACCTGGATGATGGTGGAGCCTTCAATATCAATG	1109
Qy	1031	GGAAGCACTCACTTTATTAAGATCTTGGCGTGGCTGCCAAAGAGGACTCGCGCGTGT	1099
Db	1110	GGAAGCACTCACTTTATTAAGATCTTGGCGTGGCTGCCAAAGAGGACTCTGCGCGTGT	1168
Qy	1091	GCTCTGGAAGTGCAGAAAGCAAGGTGGAGTTGGTGCCTTCAGATATTATCAGTTACAA	1156









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 12:51:18 ; Search time 136.658 Seconds  
(without alignments)  
11200.072 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_57\_623

Perfect score: 567  
Sequence: 1 CCTATCAGATTATCTTAACA.....AGATTGTGTATTAACCAA 567

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

```

1: N_Geneseq_19Jun03.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	567	100.0	1851	19	AAV82779
2	567	100.0	1851	2	ABO92016
3	567	100.0	1863	21	AAZ98034
4	567	100.0	1863	22	AAAD1647
5	567	100.0	1863	24	ABK68743
6	567	100.0	1863	25	ACC50817
7	567	100.0	1863	25	ABZ71453
8	567	100.0	1884	21	AAZ58313

9	567	100.0	1895	22	AAH99703
10	567	100.0	1923	21	AAZ98139
11	567	100.0	2077	25	ACC50603
12	567	100.0	2077	25	ABZ71331
13	563.8	99.4	1860	22	AAK94491
14	555	97.9	609	20	AAV87487
15	546.4	96.4	895	21	AAA44369
16	526.4	92.8	1767	21	AAA40493
17	526.2	92.8	1778	18	AAV02296
18	525	92.6	1784	23	AA573592
19	512.2	90.3	642	22	AAK92181
20	512.2	90.3	642	22	AAK93447
21	399.6	70.5	1778	20	ABK62598
22	290.8	51.3	314	20	AAV86369
23	284.6	50.2	462	18	AAV87398
24	284.6	50.2	462	18	AAV88060
25	284.6	50.2	462	18	AAV02139
26	277.2	48.9	441	22	AAK56831
27	254.6	44.9	317	20	AAK40587
28	121	21.3	217	20	AAK40583
29	60	10.6	60	24	ABN36212
30	49	8.6	424	23	AA573591
31	41	7.2	65	24	ABN28454
32	39.4	6.9	975	21	AACT7173
33	39.4	6.9	1556	24	AAAD3726
34	39.4	6.9	2319	25	ABK63579
35	39.4	6.9	2761	25	ABK34836
36	39.4	6.9	3383	23	ABV30133
37	39.4	6.9	8211	22	AAK80450
38	37.8	6.7	1353	24	ABN70123
39	36.8	6.5	8496	24	ABQ76621
40	36.6	6.5	766431	24	ABQ74964
41	36.2	6.4	1767	22	AAI06993
42	36	6.3	17967	22	ABJ33014
43	36	6.3	22655	22	AAK70122
44	36	6.3	2365589	24	ABA90521
45	35.4	6.2	605	22	AAK85262

#### ALIGNMENTS

RESULT 1  
ID AAV82779 standard; CDNA; 1851 BP.  
AAV82779;  
25-FEB-1999 (first entry)  
Clone b45\_2 isolated from human adult placenta CDNA library.  
Secreted protein; nutritional activity; immune stimulating; vaccine;  
suppressing activity; haematopoiesis regulating activity;  
tissue growth activity; activin; inhibin activity; chemotaxis;  
chemokine activity; haemostasis; thrombolytic activity; receptor;  
ligand; anti-inflammatory; cadherin; tumour invasion suppressor;  
tumour inhibition; gene therapy; ds.  
Homo sapiens.  
WO9842739-A2.  
01-OCT-1998.  
20-MAR-1998; 98WO-US05653.  
19-MAR-1998; 98US-004446.  
21-MAR-1997; 97US-0822167.  
(GENY) GENETICS INST INC.  
Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

Query Match	100.0%	Score 567	DB 19	Length 1851
Best Local Similarity	100.0%	Pred. No. 4.2e-161		
Matches 567	Conservative 0	Mismatches 0	Indels 0	Gaps 0
<p>Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;</p>				
Query	1	CCTATCAGATTTCCTTAACAAGAAAACCAACGTGAAAAAATAAATTCCTATCTTC	60	
Db	57	CCTATCAGATTTCCTTAACAAGAAAACCAACGTGAAAAAATAAATTCCTATCTTC	116	
Query	61	GCATTTTCGGTGAGTTCACCTTTTATCCCTGCTCTGGAAAGCTATAATGCAAGAT	120	
Db	117	GCATTTTCGGTGAGTTCACCTTTTATCCCTGCTCTGGAAAGCTATAATGCAAGAT	176	
Query	121	GGCATCTCTAAGAGGACCTTTGAAAGAAATAAAGAAATATGACAGCTGTGGAGATGT	180	
Db	177	GGCATCTCTAAGAGGACCTTTGAAAGAAATAAAGAAATATGACAGCTGTGGAGATGT	236	
Query	181	GCTAAGCAATATATCACTTACTGTGTTATGTTAAAGCCAGAAACGATCCTATAGGGA	240	
Db	237	GCTAAGCAATATATCACTTACTGTGTTATGTTAAAGCCAGAAACGATCCTATAGGGA	296	
Query	241	TTGGACATCTGCTGGATGATCTGTTGGAACCCAGACGTAGTGCTCCAGAAACCTAGAAAA	300	
Db	297	TTGGACATCTGCTGGATGATCTGTTGGAACCCAGACGTAGTGCTCCAGAAACCTAGAAAA	356	
Query	301	GGCATCCAAATTTATGTATCCAAAACTGTCAGCAAGATGGGCTTGAGAAAAATTCACTGGAG	360	
Db	357	GGCATCCAAATTTATGTATCCAAAACTGTCAGCAAGATGGGCTTGAGAAAAATTCACTGGAG	416	
Query	361	CCAGTGAAGAAATCCCCACCTGGAGAGGGGAGAAAGATCAGCTGTGATGCTGGAAGCCAGA	420	
Db	417	CCAGTGAAGAAATCCCCACCTGGAGAGGGGAGAAAGATCAGCTGTGATGCTGGAAGCCAGA	476	
Query	421	ATTCAATTAAGATAGCCATCTGGGCTCTGGCAGACAGATTGGGAGCTCCTCCAGAAAGCAAT	480	
Db	477	ATTCAATTAAGATAGCCATCTGGGCTCTGGCAGACAGATTGGGAGCTCCTCCAGAAAGCAAT	536	
Query	481	ACAGCAGAAATTCGTGTGTGTAACCTCTTTCGATGGAATGTCAGAGAAAGGCTCAGAACCA	540	
Db	537	ACAGCAGAAATTCGTGTGTGTAACCTCTTTCGATGGAATGTCAGAGAAAGGCTCAGAACCA	596	
Query	541	AGAGGAAAGATTTGTTTATACCAA 567		
Db	597	AGAGGAAAGATTTGTTTATACCAA 623		

ID	ABQ92016 standard; cDNA; 1851 BP
XX	
AC	ABQ92016;
DT	04-OCT-2002 (first entry)
DE	Human polynucleotide SEQ ID NO 13.
XX	
KW	Human; cytosolic; antirheumatic; antiarthritic; vulnery; analgesic; antihlammatory; antibacterial; immunosuppressive; antiparkinsonian; neuroprotective; nocotropic; osteopathic; haemostatic; vasotropic; anticancer; fungicide; antidiabetic; antiaschmatic; antiallergic; immunostimulant; antiparasitic; secreted protein; transmembrane protein; cytokine; cell proliferation; cell differentiation; autoimmune disease; stem cell; growth factor; nervous system disease; neuropathy; Alzheimer's disease; Parkinson's disease; Huntington's disease; Atherosclerosis; severe combined immunodeficiency; SCID; infection; multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.
OS	Homo sapiens.
XX	
PN	US2002065394-A1.
XX	
PD	30-MAY-2002.
XX	
PF	22-DEC-2000; 2000US-0745763.
XX	
PR	18-MAR-1998; 98US-0040963.
XX	
PA	(JACO/) JACOBS K. (MCCO/) MCCOY J M. (LAVA/) LAVALLIE E R. (COLL/) COLLINS-RACIE L A. (EVAN/) EVANS C. (MERB/) MERBERG D. (TREAA/) TREACY M. (SPAUL/) SPAULDING V.
XX	
F1	Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
P1	Merberg D, Treacy M, Spaulding V;
DR	WPI: 2002-582343/62.
XX	
DR	P-PDSB; ABP61800.
PT	
PT	Novel secreted or transmembrane protein and polynucleotide encoding the protein, useful for diagnosis and treatment of neurological disorders, cancer, autoimmune diseases, bone disorders and lung or liver fibrosis
PT	-
XX	
PS	Claim 50; Page 113-114; 284pp; English.
XX	
XX	The invention relates to human secreted or transmembrane protein (I), their fragments and is encoded by specific complementary deoxyribonucleic acid (cDNA) inserts (II), where the protein is substantially free from other mammalian proteins. (I) are useful for preventing, treating or ameliorating a medical condition, especially immunological treatment or prevention of tumours. (I) exhibits activity relating to angiogenesis, cytokine, cell proliferation, cell differentiation, antiinflammatory, stem cell growth factor activity and activin or inhibin-related activities. (I) can be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. (I) induces the proliferation of neural cells and regeneration of nerve and brain tissue and is useful for the treatment of central and peripheral nervous system diseases and neuropathies, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis. (I) is involved in chemotactic or chemokinetic activity, regulation of haematopoiesis and is useful for treating myeloid or lymphoid cell disorders, platelet disorders such as thrombocytopenia and for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue growth and in tissue repair, healing of burns, incisions, ulcers, for treating osteoporosis, osteoarthritis, bone degenerative disorders or periodontal disease. (I) is also useful for gut protection or

CC regeneration and treatment of lung or liver fibrosis, reperfusion injury  
 CC in various tissues, various immune deficiencies and disorders including  
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,  
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,  
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,  
 CC such as asthma or other respiratory problems. (11) is useful to express  
 CC recombinant protein, as markers for tissues in which the corresponding  
 CC protein is preferentially expressed and in gene therapy. The present  
 CC sequence is that of a polynucleotide of the invention.

XX  
 SQ Sequence 1851 BP, 531 A; 413 C; 438 G; 469 T; 0 other;

Query Match 100.0%; Score 567; DB 24; Length 1851;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-161;

Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGATTCGATTATCTTAACAGAAACCACTGGAAAAAATGAAATTCCTTATCTTC 60  
 DB 57 CCGATTCGATTATCTTAACAGAAACCACTGGAAAAAATGAAATTCCTTATCTTC 116  
 QY 61 GCATTTTGGTGGTGTTCACCTTTTACCTGCTGTGGGAAAGCTATATGCAAGAT 120  
 DB 117 GCATTTTGGTGGTGTTCACCTTTTATCCTGTGCTGTGGGAAAGCTATATGCAAGAT 176  
 QY 121 GGCATCTCTAAGAGGACTTTTGAAGAATAAAGAGAAATAGCAGCTGTGAGATGTT 180  
 DB 177 GGCATCTCTAAGAGGACTTTTGAAGAATAAAGAGAAATAGCAGCTGTGAGATGTT 236  
 QY 181 GCTAAGCAATCATCAACTGAGCTGTTTATGTTAAGCCAGAACAGATCCTATGAGCA 240  
 DB 237 GCTAAGCAATCATCAACTGAGCTGTTTATGTTAAGCCAGAACAGATCCTATGAGCA 296  
 QY 241 TTGGACATCTTGGTGTGATCTGTGACCCGACACTGAGTGTCTCCAGAACTTGAAGAAA 300  
 DB 297 TTGGACATCTTGGTGTGATCTGTGACCCGACACTGAGTGTCTCCAGAACTTGAAGAAA 356  
 QY 301 GCCATCCAAATTATGATCAAAACCTGAGCAAGATGGCTGTGAGAAAGTTCACTGGAG 360  
 DB 357 GCCATCCAAATTATGATCAAAACCTGAGCAAGATGGCTGTGAGAAAGTTCACTGGAG 416  
 QY 361 CCAGTGAAGATATCCCACTGGGAGAGGGGAGAGAAATGATGATCTGTGAGCCAGCA 420  
 DB 417 CCAGTGAAGATATCCCACTGGGAGAGGGGAGAGAAATGATGATCTGTGAGCCAGCA 476  
 QY 421 ATTCTATGAATAGGATCTCTGGTCTTGGCAGCAGCATTTGGGACTCTCTCCAGAAAGGCA 480  
 DB 477 ATTCTATGAATAGGATCTCTGGTCTTGGCAGCAGCATTTGGGACTCTCTCCAGAAAGGCA 536  
 QY 481 ACAGCAGAGATTCTGTGGTGTGACCTTTGATGAACTGACAGAGAGGCTCCAGAAAGCA 540  
 DB 537 ACAGCAGAGATTCTGTGGTGTGACCTTTGATGAACTGACAGAGAGGCTCCAGAAAGCA 596  
 QY 541 AGAGGAGAGATTGTGTTTATACCAA 567  
 DB 597 AGAGGAGAGATTGTGTTTATACCAA 623

RESULT 3  
 AA298034  
 ID AA298034 standard; cDNA; 1863 BP.

XX AA298034;

XX 09-MAY-2000 (first entry)

DE Human secreted protein encoding nucleotide sequence SEQ ID NO:28.

XX Human: secreted protein; diagnosis; cytostatic; immunosuppressive;  
 KW antiinflammatory; neurotropic; neuroprotective; antiallergic; cancer;  
 KW tumour; degenerative disorder; developmental abnormality; allergy;  
 KW foetal deficiency; blood disorder; immune system disorder; arthritis;  
 KW autoimmune disease; hepatic disease; renal disease; inflammation;  
 KW Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;

KW infection; AIDS; spinal cord injury; transplant rejection; diabetes;  
 KW asthma; sepsis; acne; psoriasis; cardiovascular disorder;  
 KW reproductive disorder; gastrointestinal disorder; respiratory disorder;  
 KW metabolic disorder; food additive; preservative; ss.

XX Homo sapiens.

XX MO200004140-A1.

XX 27-JAN-2000.

XX 14-JUL-1999; 99WO-US15849.

XX 15-JUL-1998; 98US-0092921.

XX 15-JUL-1998; 98US-0092922.

XX 15-JUL-1998; 98US-0092956.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Komatsu, Duan RD, Rosen CA, Moore PA, Shi Y;  
 PI Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE;  
 PI Muenemki M, Endress GA, Soppet DR;

XX WPI; 2000-161128/14.

XX P-PSDB; AAY87081.

PT New isolated human genes, useful for diagnosis and treatment of, e.g.  
 PT cancers, neurological or blood disorders

PS Claim 1; Page 319; 494pp; English.

CC The polynucleotide sequences given in AA298017 to AA298108 encode the  
 CC human secreted proteins given in AAY87064 to AAY8723. Human secreted  
 CC protein can have activities based on the tissues and cells the genes are  
 CC expressed in. Examples of activities include: cytostatic;  
 CC immunosuppressive; antiinflammatory; neurotropic; neuroprotective; and  
 CC polypeptides are useful for preventing, treating or ameliorating medical  
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions  
 CC can be diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new  
 CC polynucleotides. Human secreted proteins and their polynucleotides can  
 CC be used for developing products for the diagnosis or treatment of cancer,  
 CC tumours, neurodegenerative disorders, developmental abnormalities and  
 CC foetal deficiencies, blood disorders, diseases of the immune system,  
 CC autoimmune diseases, hepatic and renal disease, inflammation,  
 CC allergies, Alzheimer's disease, behavioural disorder, schizophrenia,  
 CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,  
 CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,  
 CC cardiovascular disorders, reproductive disorders, gastrointestinal  
 CC disorders, respiratory disorders and metabolic disorders. The  
 CC proteins or polynucleotides can also be used as food additives or  
 CC preservatives. The proteins are also useful for identifying their  
 CC binding partners. AA298008 to AA298016 and AAY87063 are sequence used in  
 CC the exemplification of the present invention.

XX Sequence 1863 BP, 533 A; 417 C; 443 G; 470 T; 0 other;

Query Match 100.0%; Score 567; DB 21; Length 1863;

Best Local Similarity 100.0%; Pred. No. 4.2e-161; Mismatches 0; Indels 0; Gaps 0;

Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGATTCGATTATCTTAACAGAAACCACTGGAAAAAATGAAATTCCTTATCTTC 60  
 DB 57 CCGATTCGATTATCTTAACAGAAACCACTGGAAAAAATGAAATTCCTTATCTTC 116  
 QY 61 GCATTTTGGTGGTGTTCACCTTTTACCTGCTGTGGGAAAGCTATATGCAAGAT 120  
 DB 117 GCATTTTGGTGGTGTTCACCTTTTATCCTGTGCTGTGGGAAAGCTATATGCAAGAT 176  
 QY 121 GGCATCTCTAAGAGGACTTTTGAAGAATAAAGAGAAATAGCAGCTGTGAGATGTT 180  
 DB 177 GGCATCTCTAAGAGGACTTTTGAAGAATAAAGAGAAATAGCAGCTGTGAGATGTT 236

OY 181 GCTAAAGCATCATCAACCTAGCTGTTATGTTAAAGCCAGAAAGATCCATGAGGA 240  
 Db 237 GCTAAAGCATCATCAACCTAGCTGTTATGTTAAAGCCAGAAAGATCCATGAGGA 296  
 OY 241 TTGGCACTTCTGTTGATGATCTGTTGAGCCAGACTGAGTGGCTCCAGAACCTAGAAAA 300  
 Db 297 TTGGCACTTCTGTTGATGATCTGTTGAGCCAGACTGAGTGGCTCCAGAACCTAGAAAA 356  
 OY 301 GCCATCCAAATTATGATCAACCAAACTGAGCAAGATGGCTGGAGAAAGTTCACTGGAG 360  
 Db 357 GCCATCCAAATTATGATCAACCAAACTGAGCAAGATGGCTGGAGAAAGTTCACTGGAG 416  
 OY 361 CCAGTGAAGATATCCCACTGGAGAGAGGAGAGAGATGATGATCTGGAGCCAGAA 420  
 Db 417 CCAGTGAAGATATCCCACTGGAGAGAGGAGAGAGATGATGATCTGGAGCCAGAA 476  
 OY 421 ATTCATTAAGATATGATGATCTGGAGTCTGGAGAGATGATGATCTGGAGAGGAGAT 480  
 Db 477 ATTCATTAAGATATGATGATCTGGAGTCTGGAGAGATGATGATCTGGAGAGGAGAT 536  
 OY 481 ACAGAGAAAGTTCTGTTGATGATCTGTTGATGATGATGATGATGATGATGATGATGAT 540  
 Db 537 ACAGAGAAAGTTCTGTTGATGATCTGTTGATGATGATGATGATGATGATGATGATGAT 596  
 OY 541 AGAGGAGAGATGTTGTTGTTATACCA 567  
 Db 597 AGAGGAGAGATGTTGTTGTTATACCA 623

## RESULT 4

AAD11647  
 ID AAD11647 standard; cDNA; 1863 BP.

AC AAD11647;

DT 24-SEP-2001 (first entry)

DE Human secreted protein-encoding gene 18 cDNA clone HRA035, SEQ ID NO:28.

XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma;  
 XX foetal abnormality; developmental abnormality; haematopoietic disorder;  
 XX immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 XX Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;  
 XX psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;  
 XX inflammation; neurological disorder; Alzheimer's disease; food additive;  
 XX angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;  
 XX pregnancy-related disorder; endocrine disorder; infection; wound healing;  
 XX cell culture; chemotaxis; vulnery; binding partner identification;  
 XX gene therapy; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 99..1517

XX FT /tag= a

XX FT sig\_peptide 99..170

XX FT mat\_peptide 171..1514

XX FT /tag= b

XX FT /tag= c

XX FT /product= "Mature human secreted protein"

XX WO200151504-A1.

XX 19-JUL-2001.

XX 12-JAN-2001; 2001WO-US00911.

XX 13-JAN-2000; 2000US-0482273.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y;  
 PI Latifler DM, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;  
 PI Endress GA, Muscenski M, Edner R;  
 XX WPI; 2001-425865/45.  
 DR P-PSDB; AAE06058.  
 XX Isolated nucleic acid molecule encoding a human secreted protein is  
 PT used in preventing, treating or ameliorating a medical condition -  
 PS Claim 1; Page 686; 864pp; English.

CC AAD1630-AAD11721 represent cDNAs corresponding to 71 human secreted  
 CC protein genes, and AAE06041-AAE06132 represent the proteins they encode.  
 CC AAE06133-AAE06205 represent human secreted protein fragments.  
 CC The secreted proteins and their genes are useful for preventing, treating  
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
 CC Pathological conditions can be diagnosed by determining the amount of the  
 CC new protein in a sample or by determining the presence of mutations in  
 CC the new genes. Specific uses are described for each of the 71 genes,  
 CC based on the tissues in which they are most highly expressed, and include  
 CC developing products for the diagnosis or treatment of proliferative  
 CC disorders, cancer, tumours, foetal and developmental abnormalities,  
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,  
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,  
 CC pregnancy-related disorders, endocrine disorders, and infections. The  
 CC proteins can also be used to aid wound healing and epithelial cell  
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
 CC before transplantation, for supporting cell culture of primary tissues,  
 CC to regenerate tissues, to identify their cognate ligands or binding  
 CC partners, and in chemotaxis, and can be used as a food additive or  
 CC preservative to modify storage properties. Antibodies specific for a  
 CC protein of the invention can be used in alleviating symptoms associated  
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).  
 CC The present sequence represents a human secreted protein-encoding cDNA of  
 CC the invention.

CC Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

CC Query Match 100.0%; Score 567; DB 22; Length 1863;

CC Best Local Similarity 100.0%; Pred. No. 4,2e-161; Mismatches 0; Gaps 0;

CC Matches 567; Conservative 0; Indels 0;

OY 1 CCTATCAGATTATCTTAACAAAGAAACCAACTGGAAAAAATGAATTCCTATCTTC 60  
 Db 57 CCTATCAGATTATCTTAACAAAGAAACCAACTGGAAAAAATGAATTCCTATCTTC 116  
 OY 61 GCATTTTGGGTGGTGTTCACCTTTTATCCCTGTGCTTGGGAAAGCTATATGCAAGAT 120  
 Db 117 GCATTTTGGGTGGTGTTCACCTTTTATCCCTGTGCTTGGGAAAGCTATATGCAAGAT 176  
 OY 121 GGCATCTCTTAAGAGACTTTTGAAGAAATAAAGAAATAATAGCAGCTGTGAGAGTT 180  
 Db 177 GGCATCTCTTAAGAGACTTTTGAAGAAATAAAGAAATAATAGCAGCTGTGAGAGTT 236  
 OY 181 GCTAAAGCATCATCAACCTAGCTGTTATGTTAAAGCCAGAAAGATCCATGAGCGA 240  
 Db 237 GCTAAAGCATCATCAACCTAGCTGTTATGTTAAAGCCAGAAAGATCCATGAGCGA 296  
 OY 241 TTGGCACTTCTGTTGATGATCTGTTGAGCCAGACTGAGTGGCTCCAGAACCTAGAAAA 300  
 Db 297 TTGGCACTTCTGTTGATGATCTGTTGAGCCAGACTGAGTGGCTCCAGAACCTAGAAAA 356  
 OY 301 GCCATCCAAATTATGATCAACCAAACTGAGCAAGATGGCTGGAGAAAGTTCACTGGAG 360  
 Db 357 GCCATCCAAATTATGATCAACCAAACTGAGCAAGATGGCTGGAGAAAGTTCACTGGAG 416  
 OY 361 CCAGTGAAGATATCCCACTGGAGAGAGGAGAGAGATGATGATCTGGAGCCAGAA 420

Db 417 CCACGTGAGAAATACCCACCTGGAGAGGGGAGAAATACGCTGTGATCTGTGAGCAGCAAGA 476  
Qy 421 ATTGATAGATAGACCATCTGGGTCCTGGCAGCAGCATTTGGGACCTCCCGAGAAAGGCATT 480  
Db 477 ATTGATAGATAGACCATCTGGGTCCTGGCAGCAGCATTTGGGACCTCCCGAGAAAGGCATT 536  
Qy 481 ACACGAGAAAGTTCTGTGTGTGACCTCTTTCGATGAACTGACAGAAAGGGCCTCAGAAACA 540  
Db 537 ACACGAGAAAGTTCTGTGTGTGACCTCTTTCGATGAACTGACAGAAAGGGCCTCAGAAACA 596  
Qy 541 AGAGGGAAGATTGTTGTTTATTAACCA 567  
Db 597 AGAGGGAAGATTGTTGTTTATTAACCA 623

RESULT 5

ABK69743  
ID ABK69743 standard; cDNA; 1863 BP.

AC ABK69743;  
XX  
DT 15-JUL-2002 (first entry)

DE Human secreted protein gene 18 #1.

Human; sex; gene; secreted protein; gene therapy; immunosuppressive;  
antiarrhythmic; antithrombotic; antiproliferative; cytostatic; cardiac;  
vasoactive; cerebroprotective; neurotrophic; neuroprotective; antibacterial;  
viral; fungicide; antifungal; ophthalmological; autoimmune disease; neoplasm;  
rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;  
cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;  
angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
ocular disorder; corneal infection; wound healing; skin aging;  
epithelial cell proliferation; food additive.

OS Homo sapiens.  
XX  
PN W0200226931-A2.

PD 04-APR-2002.  
XX  
PF 24-SEP-2001; 2001WO-US29871.

PR 25-SEP-2000; 2000US-234925P.  
XX  
PR 12-JAN-2001; 2001WO-US00911.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Komatsu J, Duan DR, Rosen CA, Moore PA, Shi Y;  
PI Lafleur DM, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;  
PI Erdres GA, Mucenki M, Ebner R;

DR WPI; 2002-362489/39.  
XX  
DR P-PSDB; ABG33880.

PT Novel 71 isolated secreted polypeptides and polynucleotides encoding  
PT the polypeptides, useful for treating Huntington's disease, sepsis,  
PT meningitis, thrombocytopenia, haemolytic anaemia, rheumatoid arthritis,  
PT asthma

PS Claim 1; Page 1170; 1478bp; English.

XX The invention relates to an isolated nucleic acid molecule (or its  
XX fragment, homologue complement or allelic variant) encoding a human  
XX secreted protein (and its fragment, domain, epitope, variant, secreted  
XX form and species variant). Also included are a recombinant vector  
XX comprising the nucleic acid, a recombinant host cell comprising the  
XX vector, an antibody against the secreted protein, a recombinant host cell  
XX that expresses the secreted protein and a method of identifying a binding  
XX partner of the secreted protein. The nucleic acid and protein are used to  
XX prevent, diagnose, treat or ameliorate a medical condition in e.g.  
XX humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep

CC for example autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver.  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. Many other diseases and  
CC disorders are listed in the specification. The polypeptides can also be  
CC used to aid wound healing an epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence encodes a novel human secreted protein of the invention.

SO Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;  
Query Match 100.0%; Score 567; DB 24; Length 1863;  
Best Local Similarity 100.0%; Pred. No. 4.2e-161;  
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTATCAGATTATCTTAACAAAGAAACCACTGGAAAAAATGAATTCCTATCTTC 60  
Db 57 CCTATCAGATTATCTTAACAAAGAAACCACTGGAAAAAATGAATTCCTATCTTC 116  
Qy 61 GCATTTTCGGTGTGTGACCTTTATCCCTGTGCTGTGGAAAGCTATATGCAAGAT 120  
Db 117 GCATTTTCGGTGTGTGACCTTTATCCCTGTGCTGTGGAAAGCTATATGCAAGAT 176  
Qy 121 GGCATCTCTTAAGAGACTTTTGAAGAAATTAAGAAAGAAATAGCAGCTGTGAGATGT 180  
Db 177 GGCATCTCTTAAGAGACTTTTGAAGAAATTAAGAAAGAAATAGCAGCTGTGAGATGT 236  
Qy 181 GCTTAAGCATCATCACTACGCTGTGTATGTGTAAGCCCAAGACAGATCTATGAGCGA 240  
Db 237 GCTTAAGCATCATCACTACGCTGTGTATGTGTAAGCCCAAGACAGATCTATGAGCGA 296  
Qy 241 TTGGCAGCTCTGTGTATGATCTGTGAACCACTGAGCTGAGCAAGAACTTGAAGAAA 300  
Db 297 TTGGCAGCTCTGTGTATGATCTGTGAACCACTGAGCTGAGCAAGAACTTGAAGAAA 356  
Qy 301 GCCATCCAAATTATGTACCAAAACCTGACAGAGATGGCTGAGAAAGTTCACTGGAG 360  
Db 357 GCCATCCAAATTATGTACCAAAACCTGACAGAGATGGCTGAGAAAGTTCACTGGAG 416  
Qy 361 CCAGTGAAGATACCCCACTGGGAGAGGGAGAGAAATACAGTGTGATGTGAGCCAGA 420  
Db 417 CCAGTGAAGATACCCCACTGGGAGAGGGAGAGAAATACAGTGTGATGTGAGCCAGA 476  
Qy 421 ATTGATAGATAGACCATCTGGGTCCTGGCAGCAGCATTTGGGACCTCCCGAGAAAGGCATT 480  
Db 477 ATTGATAGATAGACCATCTGGGTCCTGGCAGCAGCATTTGGGACCTCCCGAGAAAGGCATT 536  
Qy 481 ACACGAGAAAGTTCTGTGTGTGACCTCTTTCGATGAACTGACAGAAAGGGCCTCAGAAACA 540  
Db 537 ACACGAGAAAGTTCTGTGTGTGACCTCTTTCGATGAACTGACAGAAAGGGCCTCAGAAACA 596  
Qy 541 AGAGGGAAGATTGTTGTTTATTAACCA 567  
Db 597 AGAGGGAAGATTGTTGTTTATTAACCA 623

RESULT 6

ACC50817  
ID ACC50817 standard; cDNA; 1863 BP.

AC ACC50817;  
XX  
DT 12-JUN-2003 (first entry)

DE Human secreted protein coding sequence, SEQ ID 484.

XX Cardiant; antiarrhythmic; antiarrhythmoleptic; vasoactive; cytostatic;  
XX vulnerable; antiinflammatory; neurotrophic; neuroprotective;

KM antiparkinsonian; gene therapy; human; cardiovascular disorder;  
KM gene; ss.  
XX Homo sapiens.  
OS WO200295010-A2.  
XX 28-NOV-2002.  
XX 19-MAR-2002; 2002WO-US09785.  
XX 21-MAR-2001; 2001US-277340P.  
PR 19-JUL-2001; 2001US-306171P.  
PR 13-NOV-2001; 2001US-331287P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX Rosen CA, Ruben SM;  
PI WPI; 2003-129429/12.  
XX  
XX Novel human secreted proteins, useful for detecting, preventing,  
PT diagnosing, prognosticating, treating and/or ameliorating  
PT cardiovascular disorders such as arrhythmia -  
XX  
PS Claim 21; SEQ ID 484; 1881bp; English.  
XX  
XX The present invention relates to novel human secreted proteins  
CC (ABR47633-ABR48145) and their coding sequences (ACC50344-ACC50856). The  
CC proteins and their coding sequences are useful for the preparation of a  
CC diagnostic or pharmaceutical composition for diagnosing or treating a  
CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,  
CC coronary arteriosclerosis and myocardial ischaemia), neural disorders,  
CC immune system disorders, muscular disorders, reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, renal disorders,  
CC proliferative disorders and/or cancerous diseases and conditions, for  
CC wound healing and epithelial cell proliferation, to treat inflammation or  
CC infection, for treating thrombosis and arteriosclerosis, for treating or  
CC preventing neural damage which occurs in neuronal disorders or  
CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
CC disease, to enhance bone and periodontal regeneration and aid in tissue  
CC transplants or bone grafts, to prevent skin aging or hair loss, to  
CC stimulate growth and differentiation of haematopoietic cells and bone  
CC marrow cells when used in combination with other cytokines, to maintain  
CC organs before transplantation or for supporting cell culture of primary  
CC tissues, to increase or decrease differentiation or proliferation of  
CC embryonic stem cells, or to modulate mammalian characteristics or  
CC metabolism.  
CC Note: The sequenced data for this patent was published in electronic  
CC format and is available from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;  
Query Match 100.0%; Score 567; DB 25; Length 1863;  
Best Local Similarity 100.0%; Pred. No. 4,2e-161;  
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTATCAGATTATCTTAAACAAGAAAAACAACCTGAAAAAAATGAATTCCTTATCTTC 60  
DB 57 CCTATCAGATTATCTTAAACAAGAAAAACAACCTGAAAAAAATGAATTCCTTATCTTC 116  
QY 61 GCATTTTCGGTGGTTCACCTTTTATCCCTGCTCTGGGAAAGCTATATGCAAGAT 120  
DB 117 GCATTTTCGGTGGTTCACCTTTTATCCCTGCTCTGGGAAAGCTATATGCAAGAT 176  
QY 121 GGCATCTCTAAGAGAGCTTTTGAAGAATAAAGAAATATGACAGCTGGAGATGTT 180  
DB 177 GGCATCTCTAAGAGAGCTTTTGAAGAATAAAGAAATATGACAGCTGGAGATGTT 236  
QY 181 GCTAAAGCAATCATCAACCTAGCTGTTATGTTAAAGCCAGAACAGATCTCATAGCGA 240  
DB 237 GCTAAAGCAATCATCAACCTAGCTGTTATGTTAAAGCCAGAACAGATCTCATAGCGA 296

QY 241 TTGGCACTTCCTGGTTGATACGTTTGACCCAGACTGAGTGGCTTCAAGAAACCTAGAAAAA 300  
DB 297 TTGGCACTTCCTGGTTGATACGTTTGACCCAGACTGAGTGGCTTCAAGAAACCTAGAAAAA 356  
QY 301 GCCATCCAAATTATATGACCAAAACCTGACAGAAAGATGGCTGAGAAAGTTCACTTGAG 360  
DB 357 GCCATCCAAATTATATGACCAAAACCTGACAGAAAGATGGCTGAGAAAGTTCACTTGAG 416  
QY 361 CCAATGGAATATCCCACTGGAGAGAGAGAGAGAAATACAGTGTGATGCTGAGACCAAGA 420  
DB 417 CCAATGGAATATCCCACTGGAGAGAGAGAGAGAAATACAGTGTGATGCTGAGACCAAGA 476  
QY 421 ATTATAGATATAGCAATCCTGGCTGCTGGAGAGAGAGATTTGGAACTCTCCAGAAAGCAT 480  
DB 477 ATTATAGATATAGCAATCCTGGCTGCTGGAGAGAGAGATTTGGAACTCTCCAGAAAGCAT 536  
QY 481 ACAGCAGAAAGTTCTGATGATGACCTCTTTCATGATGACAGAGAAAGGCTCAGAAACA 540  
DB 537 ACAGCAGAAAGTTCTGATGATGACCTCTTTCATGATGACAGAGAAAGGCTCAGAAACA 596  
QY 541 AGAGGGAAGATTTGTTTATTAACCA 567  
DB 597 AGAGGGAAGATTTGTTTATTAACCA 623  
RESULT 7  
AB271453  
ID AB271453 standard; cDNA; 1863 BP.  
XX  
XX AB271453;  
AC  
XX  
DT 04-APR-2003 (first entry)  
XX  
XX Secreted protein-encoding gene 142 cDNA clone HRA035, SEQ ID NO:274.  
DE  
XX  
XX Human; secreted protein; digestive disorder; gastrointestinal disorder;  
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;  
KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;  
KW immune disorder; inflammation; infection; wound healing; drug screening;  
KW chromosome identification; chromosome mapping; cytostatic; gene therapy;  
KW antiinflammatory; immunosuppressive; vulnerary; chromosome 8q22.2;  
XX  
XX gene; ss.  
OS Homo sapiens.  
XX  
XX WO200276488-A1.  
PN  
XX  
PD 03-OCT-2002.  
XX  
XX 19-MAR-2002; 2002WO-US08276.  
PE  
XX  
XX 21-MAR-2001; 2001US-277340P.  
PR 19-JUL-2001; 2001US-306171P.  
PR 13-NOV-2001; 2001US-331287P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX Rosen CA, Ruben SM;  
PI WPI; 2003-029900/02.  
XX  
XX P-PSDB; ABR00274.  
DR  
XX  
XX New human secreted proteins and nucleic acids, useful for detecting,  
PT preventing, diagnosing, prognosticating, treating and/or ameliorating  
PT e.g. gastrointestinal diseases and disorders, or cancers -  
XX  
XX  
XX Claim 21; Page 918; 1216bp; English.  
CC  
CC AB271190-AB271478 represent cDNAs corresponding to 178 human secreted  
CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.  
CC AB271479-AB271540 represent human secreted protein genomic fragments. The  
CC invention also encompasses antibodies specific for the secreted proteins,

CC the use of the secreted protein in drug screening, and recombinant  
CC vectors and host cells comprising a nucleic acid of the invention. The  
CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
CC fragments specific for the secreted protein, and modulators of protein  
CC activity are useful for diagnosing, treating, ameliorating or preventing  
CC degenerative disorders. Such conditions include disorders of the mouth,  
CC oesophagus, stomach, small intestine, large intestine, liver, biliary  
CC tract and pancreas, and include cancers of these organs and tissues. The  
CC secreted proteins and their nucleic acids may also be used in the  
CC treatment of immune disorders, inflammation, infection,  
CC hyperproliferative disorders, and to promote wound healing. Nucleic acids  
CC of the invention may be used for chromosome identification, chromosome  
CC mapping, in gene therapy, for identifying individuals from minute  
CC biological samples, as hybridisation probes, and as molecular weight  
CC markers. The present sequence represents a human secreted protein-  
CC encoding cDNA clone of the invention.

**SQ** Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Query Match	100.0%	Score 567	DB 25	Length 1863
Best Local Similarity	100.0%	Pred. No.	4.2e-161	
Matches 567	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	1	CCATCAGATTATCTTAAACAAGAAAACCACTGGAAAAAATAAATTCCTATCTTC	60
Db	57	CCATCAGATTATCTTAAACAAGAAAACCACTGGAAAAAATAAATTCCTATCTTC	116
QY	61	GCAATTTTGGTGGTGGTTCACCTTTATCCCTGCTCTGGAAAAGCTATATGCAAGAT	120
Db	117	GCATTTTGGTGGTGGTTCACCTTTTATCCCTGTGCTCTGGAAAAGCTATATGCAAGAT	176
QY	121	GGCATCTCTAAGAGCATTTTGAAGAATAAAGAATAATAGCCAGCTGTGGATGTT	180
Db	177	GGCATCTCTAAGAGGCACTTTGAAGAATAAAGAATAATAGCCAGCTGTGGATGTT	236
QY	181	GCTAAAGCANTATCAACCTAGCTGTATATGGTAAAGCCAGAACAGATCCTATGACGA	240
Db	237	GCTAAAGCANTATCAACCTAGCTGTATATGGTAAAGCCAGAACAGATCCTATGACGA	296
QY	241	TTGGCATTCTGGTTGATACTGTGGACCCAGACTAGTGGCTCCAGAACCTTGAAGAA	300
Db	297	TTGGCATTCTGGTTGATACTGTGGACCCAGACTAGTGGCTCCAGAACCTTGAAGAA	356
QY	301	GCCATCCAAATTATGTACCAAAACCTGCAGCAAGTGGGCTGGAGAAAGTTCACTGGAG	360
Db	357	GCCATCCAAATTATGTACCAAAACCTGCAGCAAGTGGGCTGGAGAAAGTTCACTGGAG	416
QY	361	CCAGTAGAATATCCCACTGGAGAGGGGAGAGATAGCTGTGATGCTGGAGCCAGA	420
Db	417	CCAGTAGAATATCCCACTGGAGAGGGGAGAGATAGCTGTGATGCTGGAGCCAGA	476
QY	421	ATTCTATPAAGATAGCCATCTGGGATCTTGGACAGACATTTGGGACTCTCCAGAGGCATT	480
Db	477	ATTCTATPAAGATAGCCATCTGGGATCTTGGACAGACATTTGGGACTCTCCAGAGGCATT	536
QY	481	ACAGCAGAAAGTCTGGTGGTGAAGCTTTTCATGATGAACTGCAGAGAGGGCTCAGAAACA	540
Db	537	ACAGCAGAAAGTCTGGTGGTGAAGCTTTTCATGATGAACTGCAGAGAGGGCTCAGAAACA	596
QY	541	AGAGGGAAGATTGTGTTTATTAACCA	567
Db	597	AGAGGGAAGATTGTGTTTATTAACCA	623

RESULT 8	
AAZ58313	
ID	AAZ58313 standard; cDNA; 1884 BP.
XX	
AC	AAZ58313;
XX	
DT	08-MAY-2000 (first entry)
XX	
DE	Human peptidase NAA1AD-ase IV cDNA

XX NAAAD-ase IV; N-acetylated alpha-linked acidic dipeptidase; human  
KM chromosome 8q21.3; prostate cancer; neurodegenerative disease;  
KM Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;  
KM peripheral neuropathy; Huntington's disease; acute brain injury;  
KM multiple sclerosis; peripheral nerve trauma; ischaemia; dementia;  
KM gene therapy; diagnosis; nootropic; neuroprotective; neuroleptic;  
KM antiparkinsonian; anticonvulsant; vasotropic; ss.

OS Homo sapiens

FH	Key	Location/Qualifiers
EM	CDC	140 1557

```
/*tag= a
ET
YY
```

PN WO200004157-A2  
YY

PD 27-JAN-2000  
YY

PF 14-JUL-1999; 99WO-GB02241.  
VY

PR 14-JUL-1998; 98GB-0015284.  
VY

PA (JANC ) JANSSEN PHARM NV.  
VY

PI Pangalos M, Neefs JEFM, Peeters DCG;  
 XY

DR WPI; 2000-182424/16.  
D-DEDA 2AYE0070

Now, here's a note:

PT neural disorders  
PT Parkinson's disease

[illegible]

Pho report comments for list of DNA codes for human

CC N-acetylated alpha-linked acidic dipeptidase IV (NAALAD-ase IV, CC see AAY58879) The cDNA was obtained from a cDNA bladder cDNA

CC library: Analysis of the open reading frame predicts a type II integral membrane protein with 5 potential N-glycosylation sites

CC The NAALAD-ase II gene was mapped to chromosome 8q21.3.

CC The invention provides human NAALAD-ase L, II and IV polypeptides,  
CC CDNA antisense nucleic acid vectors, host cells, transgenic

CC organisms, antagonists and agonists. These are useful for treating  
CC novel disorders such as Alzheimer's disease, schizophrenia, AIDS

CC Parkinson's disease, peripheral neuropathy, Huntington's disease,  
CC acute brain injury, multiple sclerosis, exposure to neurotoxins

peripheral nerve trauma, ischaemia or dementia (claimed). Nucleic acids can also be used for gene therapy and for genetic screening.

CC of predisposition to disorders associated with NAALAD-ase.  
XY

Sequence 1884 BP; 520 A; 432 C; 458 G; 474 T; 0 other;

Query Match	100.0%;	Score 567;	DB 21;	Length 1884;
Best Local Similarity	100.0%;	Pred. No. 4	2e-161;	
Matches 567;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

Oy      1 CCATCAGATTATCTTTAACAGAAACCACCTGGAAAAAAATGAAATTCTTATCTTC 60
        |||||
Db      107 CCTATCAGATTATCTTTAACAGAAACCACCTGGAAAAAAATGAAATTCTTATCTTC 160

```

61 GCATTTTTCGGTGTTCACCTTTATCCCTGTGCTCTGGGAAGCTATATGCAAGAT 120

Db 167 GCATTTTCGGTGTGTTCACTTTTATCCCTGTGCTCTGGAAAGCTATATGCAAGAT 226

121 GGCATCTCTAAGAGGACTTTTGAGGAATAAAGAGAATAAGCCAGCTGTGAGATGTT 180

Db 227 GGCATCTCTAAGAGGACTTTTGAAGAAATAAAGAGAATAGCCAGCTGTGGAGATGTT 286

181 GCTAAGCAATCATCAACCTAGCTGTTATGTGTAAGCCAGAACAGATCCTATGAGCGA 240

.....



Db 287 GCTAAAGCAATCATCAACCTAGTGTATTATGTAAAGCCAGAAACAGATCCTATGAGCA 346  
 QY 241 TTGGCACTTCCTGGTGTATCTAGTGTGACCCGACGATGCTGCTCCAGAAACCTTGAAGAA 300  
 Db 347 TTGGCACTTCCTGGTGTATCTAGTGTGACCCGACGATGCTGCTCCAGAAACCTTGAAGAA 406  
 QY 301 GCCATCCAAATATATGATCAAAACCTGACGAGATGGGCTGAGAAAGTTCCACTGGAG 360  
 Db 407 GCCATCCAAATATATGATCAAAACCTGACGAGATGGGCTGAGAAAGTTCCACTGGAG 466  
 QY 361 CCAGTGAAGATACCCCACTGGAGAGGGGAGAGAAATCAGCTGTGATGCTGAGCCAGCA 420  
 Db 467 CCAGTGAAGATACCCCACTGGAGAGGGGAGAGAAATCAGCTGTGATGCTGAGCCAGCA 526  
 QY 421 ATTATATAGATAGCAATCCCTGGTCTTGGGACGACATTTGGGATCTCTCCAGAAAGCAT 480  
 Db 527 ATTATATAGATAGCAATCCCTGGTCTTGGGACGACATTTGGGATCTCTCCAGAAAGCAT 586  
 QY 481 ACAGCAGAAAGTTCTGGTGTGACCTCTTTCGATGAACTGACAGAGAGGCTCAGAAAGCA 540  
 Db 587 ACAGCAGAAAGTTCTGGTGTGACCTCTTTCGATGAACTGACAGAGAGGCTCAGAAAGCA 646  
 QY 541 AGAGGGAAGATTTGTTTATTAACCAA 567  
 Db 647 AGAGGGAAGATTTGTTTATTAACCAA 673

## RESULT 9

AAH9703 standard; cDNA, 1895 BP.

AAH9703;

16-OCT-2001 (first entry)

Human protein encoding cDNA sequence SEQ ID NO:538.

Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
 anti-inflammation; antirheumatic; antiarthritic; immunosuppressive;  
 anti-bacterial; endocrine; cardiac; central nervous system; vitruide;  
 anti-HIV; fungicide; antimutagen; cardiovascular; anti-anemic; anemia;  
 antiaggregant; haemostatic; vulnary; anti-ulcer; osteopathic; eczema;  
 dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic;  
 neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
 immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
 anti-anaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
 cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
 genetic disease; haematopoietic disorder; platelet disorder; asthma;  
 thrombocytopenia; osteoporosis; severe combined immunodeficiency;  
 allergic rhinitis; diabetes; multiple sclerosis; depression;  
 Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
 neurological disorder; ss.

Homo sapiens.

WO200153455-A2.

26-JUL-2001.

22-DEC-2000; 2000WO-US35017.

23-DEC-1999; 99US-0471275.

21-JAN-2000; 2000US-0488725.

25-APR-2000; 2000US-0552317.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Dymnac RT;

WPI, 2001-457603/49.

P-PSDB; AAM25762.

Isolated human polynucleotides encoding polypeptides, useful for the

PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -

XX Claim 1; Page 591; 1217pp; English.

CC AAH9166 to AAH9904 encode the human proteins given in AAM25225 to  
 CC AAM25963. The proteins can have activities based on the tissues and  
 CC cells they are expressed in, such as: anti-inflammatory; antirheumatic;  
 CC antiallergic; immunosuppressive; antibacterial; endocrine; cardiac;  
 CC central nervous system; vitruide; anti-HIV; fungicide; antimutagen;  
 CC cardiovascular; anti-anemic; antiaagregant; haemostatic; vulnary;  
 CC antidiabetic; osteopathic; dermatological; antiallergic; antiseptic;  
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
 CC encoding them can be used in gene therapy, antisense therapy and vaccine  
 CC production. The proteins and polynucleotides are useful for screening for  
 CC agonists or antagonists of a protein and for the treatment and diagnosis  
 CC of disorders associated with the activity of a protein e.g. inflammation,  
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,  
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
 CC neurological disorders.

SO Sequence 1895 BP; 530 A; 439 C; 450 G; 476 T; 0 other;

Query Match 100.0%; Score 567; DB 22; Length 1895;

Best Local Similarity 100.0%; Pred. No. 4,2e-161; Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTATCAGATTAATCTTAACAGAGAAACCACTGGAAAAAATGAATTCCTATTATTC 60  
 Db 84 CCTATCAGATTAATCTTAACAGAGAAACCACTGGAAAAAATGAATTCCTATTATTC 143  
 QY 61 GCATTTTCGGTGTGTCACCTTTATCCCTGCTCTGGAAAGCTATATCAAGAAAT 120  
 Db 144 GCATTTTCGGTGTGTCACCTTTATCCCTGCTCTGGAAAGCTATATCAAGAAAT 203  
 QY 121 GGCATCTTAAGAGACTTTTGAAGAAATTAAGAAATATAGCCAGCTGTGAGATGT 180  
 Db 204 GGCATCTTAAGAGACTTTTGAAGAAATTAAGAAATATAGCCAGCTGTGAGATGT 263  
 QY 181 GCTAAAGCAATCATCAACCTAGCTGTTTATGTAAAGCCAGAAACAGATCTTATGAGCA 240  
 Db 264 GCTAAAGCAATCATCAACCTAGCTGTTTATGTAAAGCCAGAAACAGATCTTATGAGCA 323  
 QY 241 TTGGCACTTCCTGGTGTATCTGTTGGACCCAGACTGAGTGTCCAGAACTAGAAAAA 300  
 Db 324 TTGGCACTTCCTGGTGTATCTGTTGGACCCAGACTGAGTGTCCAGAACTAGAAAAA 383  
 QY 301 GCCATCCAAATTAATGATCAAAACCTGACAGAAAGATGGCTGAGAAAGTTCACTGGAG 360  
 Db 384 GCCATCCAAATTAATGATCAAAACCTGACAGAAAGATGGCTGAGAAAGTTCACTGGAG 443  
 QY 361 CCAGTGAAGATACCCCACTGGAGAGGGGAGAGAAATCAGCTGTGATGCTGAGCCAGCA 420  
 Db 444 CCAGTGAAGATACCCCACTGGAGAGGGGAGAGAAATCAGCTGTGATGCTGAGCCAGCA 503  
 QY 421 ATTATATAGATAGCAATCCCTGGTCTTGGGACGAGCATTTGGGATCTCTCAGAAAGCAT 480  
 Db 504 ATTATATAGATAGCAATCCCTGGTCTTGGGACGAGCATTTGGGATCTCTCAGAAAGCAT 563  
 QY 481 ACAGCAGAAAGTTCTGGTGTGACCTCTTTCGATGAACTGACAGAGAGGCTCAGAAAGCA 540  
 Db 564 ACAGCAGAAAGTTCTGGTGTGACCTCTTTCGATGAACTGACAGAGAGGCTCAGAAAGCA 623  
 QY 541 AGAGGGAAGATTTGTTTATTAACCAA 567  
 Db 624 AGAGGGAAGATTTGTTTATTAACCAA 650



RESULT 10  
AAZ98139  
ID AAZ98139 standard; cDNA; 1923 BP.  
XX  
AC AAZ98139;  
XX  
DT 11-MAY-2000 (first entry)  
XX  
DE Human signal peptide containing protein HSP-31 cDNA SEQ ID NO:165.  
XX  
KW Human, signal peptide-containing protein; HSP; diagnosis; cancer;  
KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;  
KW anti-infective; neuroprotective; cardiovascular; hepatotropic;  
KW antiaesthetic; gene therapy; cell proliferation; neurological disorder;  
KW reproductive disorder; developmental disorder; arteriosclerosis;  
KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;  
KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;  
KW Parkinson's disease; Huntington's disease; ovulatory defect;  
KW muscular dystrophy; ss.  
XX  
OS Homo sapiens.  
XX  
PN W0200000610-A2.  
XX  
PD 06-JAN-2000.  
XX  
PF 25-JUN-1999; 99WO-US14484.  
XX  
PR 26-JUN-1998; 98US-0090762.  
PR 31-JUL-1998; 98US-0094983.  
PR 01-OCT-1998; 98US-0102686.  
PR 11-DEC-1998; 98US-0112129.  
XX  
PA (INCY-) INCYTE PHARM INC.  
XX  
PI Lei P, Tang YT, Gorgone GA, Corley NC, Guegler KU, Baughn MR;  
PI Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;  
PI Bandman O;  
XX  
DR WPI: 2000-150673/14.  
DR P-PSDB; AAY87254.  
XX  
PT New human signal peptide-containing proteins useful in treatment,  
PT prevention and diagnosis of e.g. cancer, inflammation and  
PT cardiovascular disease -  
XX  
PS Claim 9; Page 269-270; 327pp; English.  
XX  
XX AAZ98109 to AAZ98242 encode AAY87224 to AAY87357 which represent the  
CC human signal peptide-containing proteins HSP-1 to HSP-134. HSPs have  
CC anticancer, anti-inflammatory, antimicrobial, nootropic, hepatotropic,  
CC neuroprotective, cardiovascular and antiaesthetic activities, and can  
CC be used in gene therapy. HSPs can be used to treat or prevent disorders  
CC associated with decreased activity or function of HSP. Antagonists of  
CC HSP are used to treat or prevent disorders associated with increased  
CC activity or function of HSP. Such diseases include cell proliferation  
CC (including cancer), inflammation, cardiovascular, neurological,  
CC reproductive or developmental disorders, (e.g. arteriosclerosis,  
CC cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,  
CC asthma, Crohn's disease, microbial or other infections, congestive or  
CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's  
CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSP  
CC nucleic acids can be used for the recombinant production of HSP, for  
CC detecting HSP in standard hybridisation and amplification assays (for  
CC diagnosis and monitoring), in gene therapy, as antisense,  
CC triplex-forming or ribozyme therapeutics, for detecting related sequences  
CC or genetic variations, and for chromosomal mapping. HSP are also used to  
CC raise specific antibodies (Ab) and to screen for agonists and  
CC antagonists (potential therapeutic agents). Ab are used to diagnose, or  
CC monitor, HSP-related diseases (in usual immunoassays), as therapeutic  
CC antagonists, in competitive drug screens, and for purification of HSP  
CC from natural sources.  
XX

SQ Sequence 1923 BP; 538 A; 439 C; 471 G; 475 T; 0 other;  
Query Match 100.0%; Score 567; DB 21; Length 1923;  
Best Local Similarity 100.0%; Pred. No. 4.3e-161;  
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTATCAGATTATCTTAACAAGAAACCAACCTGGAAAAAATGAAATTCCTATCTTC 60  
DB 137 CCTATCAGATTATCTTAACAAGAAACCAACCTGGAAAAAATGAAATTCCTATCTTC 196  
QY 61 GCATTTTCGGTGTGTTGACCTTTATCCCTGTCTGTGGAAAGCTATATGCAAGAT 120  
DB 197 GCATTTTCGGTGTGTTGACCTTTATCCCTGTCTGTGGAAAGCTATATGCAAGAT 256  
QY 121 GGCACTCTTAAGAGGACTTTTGAAGAAATTAAGAAATATGACCTGTGGAGATGT 180  
DB 257 GGCACTCTTAAGAGGACTTTTGAAGAAATTAAGAAATATGACCTGTGGAGATGT 316  
QY 181 GCTAAAGCAATCATCACTAGCTTTATGTATGTAAGCCAGAAACAGATCCTATGACGA 240  
DB 317 GCTAAAGCAATCATCACTAGCTTTATGTATGTAAGCCAGAAACAGATCCTATGACGA 376  
QY 241 TTGGCACTTCTGTTGATPACTGTTTGAACCAAGTGTGCTTCAAGAACTTGAAGAA 300  
DB 377 TTGGCACTTCTGTTGATPACTGTTTGAACCAAGTGTGCTTCAAGAACTTGAAGAA 436  
QY 301 GCCATCCAAATTAATGTAACCAAAACCTGCAAGATGAGGCTGGAAGAAATTCCTGAG 360  
DB 437 GCCATCCAAATTAATGTAACCAAAACCTGCAAGATGAGGCTGGAAGAAATTCCTGAG 496  
QY 361 CCAATGAGAAATACCCCACTGGAGAGAGGAGAAAGATCAGCTGTGATGCTGAGCCAGA 420  
DB 497 CCAATGAGAAATACCCCACTGGAGAGAGGAGAAAGATCAGCTGTGATGCTGAGCCAGA 556  
QY 421 ATTCAATAGATAGGCACTCTGCGTCTTGGACAGCAGATTGGAGATCTCTCAAGAGCAT 480  
DB 557 ATTCAATAGATAGGCACTCTGCGTCTTGGACAGCAGATTGGAGATCTCTCAAGAGCAT 616  
QY 481 ACAGCAGAAAGTTCGGTGTGATGACCTCTTGGATGAACTGCAAGAAAGGCTCGAAGCA 540  
DB 617 ACAGCAGAAAGTTCGGTGTGATGACCTCTTGGATGAACTGCAAGAAAGGCTCGAAGCA 676  
QY 541 AGAGGAAAGATTGTTGTTTAAACCAA 567  
DB 677 AGAGGAAAGATTGTTGTTTAAACCAA 703  
RESULT 11  
ACCS0603  
ID ACCS0603 standard; cDNA; 2077 BP.  
XX  
AC ACCS0603;  
XX  
DT 12-JUN-2003 (first entry)  
XX  
DE Human secreted protein coding sequence, SEQ ID 270.  
XX  
KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cyostatic;  
KW vulnerability; antiinflammatory; nootropic; neuroprotective;  
KW antiparkinsonian; gene therapy; human; cardiovascular disorder;  
KW gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN W0200295010-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 19-MAR-2002; 2002WO-US09785.  
XX  
PR 21-MAR-2001; 2001US-277340P.  
PR 19-JUL-2001; 2001US-306171P.  
PR 13-NOV-2001; 2001US-331287P.  
XX

XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Rosen CA, Ruben SM;  
 PI WPI, 2003-129429/12.  
 DR Novel human secreted proteins, useful for detecting, preventing,  
 XX diagnosing, prognosticating, treating and/or ameliorating  
 PT cardiovascular disorders such as arrhythmia -  
 XX Claim 21, SEQ ID 270, 188bp; English.  
 PS  
 XX The present invention relates to novel human secreted proteins  
 CC (ABR47633-ABR48145) and their coding sequences (ACC50344-ACC50866). The  
 CC proteins and their coding sequences are useful for the preparation of a  
 CC diagnostic or pharmaceutical composition for diagnosing or treating a  
 CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,  
 CC coronary arteriosclerosis and myocardial ischaemia), neural disorders,  
 CC immune system disorders, muscular disorders, reproductive disorders,  
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,  
 CC proliferative disorders and/or cancerous diseases and conditions, for  
 CC wound healing and epithelial cell proliferation, to treat inflammation or  
 CC infection, for treating thrombosis and arteriosclerosis, for treating or  
 CC preventing neural damage which occurs in Alzheimer's disease and Parkinson's  
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
 CC disease, to enhance bone and periodontal regeneration and aid in tissue  
 CC transplants or bone grafts, to prevent skin aging or hair loss, to  
 CC stimulate growth and differentiation of haematopoietic cells and bone  
 CC marrow cells when used in combination with other cytokines, to maintain  
 CC organs before transplantation or for supporting cell culture of primary  
 CC tissues, to increase or decrease differentiation or proliferation of  
 CC embryonic stem cells, or to modulate mammalian characteristics or  
 CC metabolism.  
 CC Note: The sequence data for this patent was published in electronic  
 CC format and is available from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX Sequence 2077 BP; 597 A; 436 C; 482 G; 562 T; 0 other;  
 SO  
 Query Match 100.0%; Score 567; DB 25; Length 2077;  
 Best Local Similarity 100.0%; Pred. No. 4,4e-161;  
 Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 CCTATCAGATTATCTTAACAAGAAAACCACTGAGAAAAAATAAATCTCTTATCTTC 60  
 Db 90 CCTATCAGATTATCTTAACAAGAAAACCACTGAGAAAAAATAAATCTCTTATCTTC 149  
 Oy 61 GCATTTTCGGTGGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120  
 Db 150 GCATTTTCGGTGGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 209  
 Oy 121 GGCATCTCTTAAGAGACTTTTGAAGAATAAAGAAATATGCCAGCTGTGAGATGTT 180  
 Db 210 GGCATCTCTTAAGAGACTTTTGAAGAATAAAGAAATATGCCAGCTGTGAGATGTT 269  
 Oy 181 GCTAAAGCAATCATGACCTGCTTTATGTTAAGCCGAGACAGATCCATGAGGCA 240  
 Db 270 GCTAAAGCAATCATGACCTGCTTTATGTTAAGCCGAGACAGATCCATGAGGCA 329  
 Oy 241 TTGGCACTTCTGTTGATATCTTGGACCCAGACTGATGCTGCCAAGACCTAGAAAAA 300  
 Db 330 TTGGCACTTCTGTTGATATCTTGGACCCAGACTGATGCTGCCAAGACCTAGAAAAA 389  
 Oy 301 GCCATCCAAATTATATACCAAAACCTGCAGCAAGATGAGGCTGAGAAAGTTCACTGGAG 360  
 Db 390 GCCATCCAAATTATATACCAAAACCTGCAGCAAGATGAGGCTGAGAAAGTTCACTGGAG 449  
 Oy 361 CCAGGGAATATATCCCACTGGAGAGGGGAGAGAAATAGCTGTGATCTGAGAGCAAGA 420  
 Db 450 CCAGGGAATATATCCCACTGGAGAGGGGAGAGAAATAGCTGTGATCTGAGAGCAAGA 509  
 Oy 421 ATTCAATAGATATGACATCTGGGTCTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCAT 480

Db 510 ATTCAATAGATATGACATCTGGGTCTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCAT 569  
 Oy 481 ACAGCAGAGATTCTGCGTGTGACCTCTTTGATGAATGCAAGAGAGGCTCAGAAACA 540  
 Db 570 ACAGCAGAGATTCTGCGTGTGACCTCTTTGATGAATGCAAGAGAGGCTCAGAAACA 629  
 Oy 541 AGAGGGAAGATTGTTGTTTATTAACCA 567  
 Db 630 AGAGGGAAGATTGTTGTTTATTAACCA 656  
 RESULT 12  
 ID ABR71331 standard; cDNA; 2077 BP.  
 XX  
 AC ABR71331;  
 XX  
 DT 04-APR-2003 (first entry)  
 XX  
 DE Secreted protein-encoding gene 142 cDNA clone HRACJ35, SEQ ID NO:152.  
 XX  
 KW Human; secreted protein; digestive disorder; gastrointestinal disorder;  
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;  
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;  
 KW immune disorder; inflammation; infection; wound healing; drug screening;  
 KW chromosome identification; chromosome mapping; cytogenetic; gene therapy;  
 KW antiinflammatory; immunosuppressive; vulnerary; chromosome 8q22.2;  
 KW gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN MO200276488-A1.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 19-MAR-2002; 2002MO-US08276.  
 XX  
 PR 21-MAR-2001; 2001US-277340P.  
 PR 19-JUL-2001; 2001US-306171P.  
 PR 13-NOV-2001; 2001US-331287P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen CA, Ruben SM;  
 DR WPI, 2003-029900/02.  
 DR P-PSDB; ABR00152.  
 XX  
 PT New human secreted proteins and nucleic acids, useful for detecting,  
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating,  
 PT e.g. gastrointestinal diseases and disorders, or cancers -  
 XX  
 XX Claim 21; Page 841-842; 1216pp; English.  
 PS  
 AB271190-AB271478 represent cDNAs corresponding to 178 human secreted  
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.  
 CC ABR71479-ABR71540 represent human secreted protein genomic fragments. The  
 CC invention also encompasses antibodies specific for the secreted proteins,  
 CC the use of the secreted proteins in drug screening, and recombinant  
 CC vectors and host cells comprising a nucleic acid of the invention. The  
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
 CC fragments specific for the secreted proteins, and modulators of protein  
 CC activity are useful for diagnosing, treating, ameliorating or preventing  
 CC digestive disorders. Such conditions include disorders of the mouth,  
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary  
 CC tract and pancreas, and include cancers of these organs and tissues. The  
 CC secreted proteins and their nucleic acids may also be used in the  
 CC treatment of immune disorders, inflammation, infection,  
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids  
 CC of the invention may be used for chromosome identification, chromosome  
 CC mapping, in gene therapy, for identifying individuals from minute  
 CC biological samples, as hybridisation probes, and as molecular weight

CC markers. The present sequence represents a human secreted protein-  
CC encoding cDNA clone of the invention.

XX Sequence 2077 BP; 597 A; 436 C; 482 G; 562 T; 0 other;

Query Match 100.0%; Score 567; DB 25; Length 2077;

Best Local Similarity 100.0%; Pred. No. 4,4e-161;  
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CCTATCGATTATCTTAAACAAGAAACCACTGGAAAAAATGAAATTCCTTATCTTC 60
D 90 CTTATCGATTATCTTAAACAAGAAACCACTGGAAAAAATGAAATTCCTTATCTTC 149
QY 61 GCATTTTCGGTGTGTGTCACCTTTTATCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
D 150 GCATTTTCGGTGTGTGTCACCTTTTATCCTGTGCTCTGGGAAAGCTATATGCAAGAT 209
QY 121 GGCATCTCTAAGAGAGCTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGT 180
D 210 GGCATCTCTAAGAGAGCTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGT 269
QY 181 GCTAAGCAATCATCAACCTAGCTGTTTATAGTAAAGCCCAAGAACATCTTATAGCGA 240
D 270 GCTAAGCAATCATCAACCTAGCTGTTTATAGTAAAGCCCAAGAACATCTTATAGCGA 329
QY 241 TTGGCACTTCTGTGTGATAGTGTGAGCCAGACTGAGTGGCTCCAGAACCTTGAAGAAA 300
D 330 TTGGCACTTCTGTGTGATAGTGTGAGCCAGACTGAGTGGCTCCAGAACCTTGAAGAAA 389
QY 301 GCCATCCAAATTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGGAG 360
D 390 GCCATCCAAATTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGGAG 449
QY 361 CCAATGAGAAATACCCCACTGGGAGAGAGGAGAAAGATAGCTGTGATCTGAGCCAGA 420
D 450 CCAATGAGAAATACCCCACTGGGAGAGAGGAGAAAGATAGCTGTGATCTGAGCCAGA 509
QY 421 ATTCAATAGATAGGCAATCTGGGTCTTTGGCAGACGATTTGGGACTCCCTCAGAAAGGCA 480
D 510 ATTCAATAGATAGGCAATCTGGGTCTTTGGCAGACGATTTGGGACTCCCTCAGAAAGGCA 569
QY 481 ACACAGAAAGTCTGTGTGATAGCTCTTTCGATGAATGACAGAGAGGGCTCAGAAAGCA 540
D 570 ACACAGAAAGTCTGTGTGATAGCTCTTTCGATGAATGACAGAGAGGGCTCAGAAAGCA 629
QY 541 AGAGGGAAGATTGTTGTTTATTAACCA 567
D 630 AGAGGGAAGATTGTTGTTTATTAACCA 656
```

## RESULT 13

AAK94491  
ID AAK94491 standard; cDNA; 1860 BP.

XX AAK94491;

DT 06-NOV-2001 (first entry)

XX Human full-length cDNA, SEQ ID NO: 3328.

KM Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX Homo sapiens.

PN EP1130094-A2.

XX 05-SEP-2001.

XX 07-JUL-2000; 2000EP-0114089.

XX 08-JUL-1999; 99JP-0194486.

PR 11-JAN-2000; 2000JP-0118774.

PR 02-MAY-2000; 2000JP-0183765.

XX (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

DR WPI; 2001-524255/58.

DR P-PSDB; AAM93559.

PT 830 Primers useful for synthesizing full length cDNA clones and their  
PT use in genetic manipulation -

PS Claim 8; SEQ ID NO 3328; 1380bp + sequence listing; English.

CC The invention relates to primers for synthesizing full length cDNA  
CC clones. 830 cDNA molecules encoding a human protein have been  
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
CC molecules have been determined. Primers for synthesizing the full length  
CC cDNA are useful for clarifying the function of the protein encoded by  
CC the cDNA. The full length clones were obtained by construction of full  
CC length enriched cDNA libraries that were synthesised by the oligo-capping  
CC method. The primers enable the production of the full length cDNA easily  
CC without any special methods. The present sequence is a full length  
CC human cDNA of the invention.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 1860 BP; 515 A; 423 C; 451 G; 471 T; 0 other;

Query Match 99.4%; Score 563.8; DB 22; Length 1860;

Best Local Similarity 99.6%; Pred. No. 3.9e-160;  
Matches 565; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 CTTATCGATTATCTTAAACAAGAAACCACTGGAAAAAATGAAATTCCTTATCTTC 60
D 87 CTTATCGATTATCTTAAACAAGAAACCACTGGAAAAAATGAAATTCCTTATCTTC 146
QY 61 GCATTTTCGGTGTGTGTCACCTTTTATCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
D 147 GCATTTTCGGTGTGTGTCACCTTTTATCCTGTGCTCTGGGAAAGCTATATGCAAGAT 206
QY 121 GGCATCTCTAAGAGAGCTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGT 180
D 207 GGCATCTCTAAGAGAGCTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGT 266
QY 181 GCTAAGCAATCATCAACCTAGCTGTTTATAGTAAAGCCCAAGAACATCTTATAGCGA 240
D 267 GCTAAGCAATCATCAACCTAGCTGTTTATAGTAAAGCCCAAGAACATCTTATAGCGA 326
QY 241 TTGGCACTTCTGTGTGATAGTGTGAGCCAGACTGAGTGGCTCCAGAACCTTGAAGAAA 300
D 327 TTGGCACTTCTGTGTGATAGTGTGAGCCAGACTGAGTGGCTCCAGAACCTTGAAGAAA 386
QY 301 GCCATCCAAATTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGGAG 360
D 387 GCCATCCAAATTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGGAG 446
QY 361 CCAATGAGAAATACCCCACTGGGAGAGAGGAGAAAGATAGCTGTGATCTGAGCCAGA 420
D 447 CCAATGAGAAATACCCCACTGGGAGAGAGGAGAAAGATAGCTGTGATCTGAGCCAGA 506
QY 421 ATTCAATAGATAGGCAATCTGGGTCTTTGGCAGACGATTTGGGACTCCCTCAGAAAGGCA 480
D 507 ATTCAATAGATAGGCAATCTGGGTCTTTGGCAGACGATTTGGGACTCCCTCAGAAAGGCA 566
QY 481 ACACAGAAAGTCTGTGTGATAGCTCTTTCGATGAATGACAGAGAGGGCTCAGAAAGCA 540
D 567 ACACAGAAAGTCTGTGTGATAGCTCTTTCGATGAATGACAGAGAGGGCTCAGAAAGCA 626
QY 541 AGAGGGAAGATTGTTGTTTATTAACCA 567
D 627 AGAGGGAAGATTGTTGTTTATTAACCA 653
```

RESULT 14  
AAV87487 standard; cDNA, 609 BP.  
XX AAV87487;  
XX  
XX AAV87487;  
XX  
XX 27-Apr-1999 (first entry)  
XX  
XX EST clone BU45.  
XX  
XX Expressed sequence tag; secreted protein; haematopoiesis regulator;  
XX tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;  
XX chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;  
XX receptor; ligand; anti-inflammatory; tumour inhibitor; de.  
XX  
XX Homo sapiens.  
XX  
XX MO9845435-A2.  
XX  
XX 15-Oct-1998.  
XX  
XX 10-Apr-1998; 98WO-US06954.  
XX  
XX 10-Apr-1997; 97US-0835913.  
XX  
XX (GEMV ) GENETICS INST INC.  
XX  
XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;  
XX Racie LA, Spaulding V, Treacy M,  
XX  
XX WPI; 1999-070076/06.  
XX  
XX New polynucleotides encoding human secreted proteins - derived from  
XX e.g. human blood, kidney, foetal lung, placenta, testes, brain,  
XX ovary, pituitary, retina and colon cDNA libraries  
XX  
XX Claim 1; Page 580; 633pp; English.  
XX  
XX This sequence represents an expressed sequence tag (EST), and is a  
XX polynucleotide of the invention. The polynucleotides of the invention are  
XX all secreted EST sequences isolated from a variety of human tissue  
XX sources. The EST sequences and proteins encoded by them are predicted to  
XX have useful biological activities which would make them suitable for  
XX treating, preventing or ameliorating medical conditions in humans and  
XX animals, although no supporting data is given. Suggested activities  
XX include nutritional activity, immune stimulating or suppressing activity,  
XX haematopoiesis regulating activity, tissue growth activity,  
XX activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
XX and thrombolytic activity, receptor/ligand activity, anti-inflammatory  
XX activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
XX activity. The EST sequences are also stated to be useful for gene  
XX therapy.  
XX  
XX Sequence 609 BP; 190 A; 127 C; 148 G; 144 T; 0 other;  
XX  
XX Query Match 97.9%; Score 555; DB 20; Length 609;  
XX Best Local Similarity 99.8%; Pred. No. 1.1e-157;  
XX Matches 566; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 209 GCTAAAGCAATCATCAACCTAGCTGTTATGTATTAAGCCAGAAAGATCCTATGACGCA 268  
Qy 241 TTGGCACTTCTGTGTGATCTGTTGACCCAGACTGATGGCTCCAGAACTAGAAAA 300  
Db 269 TTGGCACTTCTGTGTGATCTGTTGACCCAGACTGATGGCTCCAGAACTAGAAAA 327  
Qy 301 GCCATCCAAATTAATGATCAAAACCTGACAGAAAGATGGCTGAGAAAGTTCACTGAG 360  
Db 328 GCCATCCAAATTAATGATCAAAACCTGACAGAAAGATGGCTGAGAAAGTTCACTGAG 387  
Qy 361 CCACTGAAATATCCCACTGAGAGAGAGAGAGAAAGATCACTGTGATCTGAGCCAAAGA 420  
Db 388 CCACTGAAATATCCCACTGAGAGAGAGAGAGAAAGATCACTGTGATCTGAGCCAAAGA 447  
Qy 421 ATTCATAGATAGCAATCCGCTGCTGAGAGAGAGATGGGACTCCTCCAGAAAGGCAAT 480  
Db 448 ATTCATAGATAGCAATCCGCTGCTGAGAGAGAGATGGGACTCCTCCAGAAAGGCAAT 507  
Qy 481 ACAGCAAGATGCTGATGATGACCTCTTTCATGATGATGAGAGAGAGGCTCAGAAACA 540  
Db 508 ACAGCAAGATGCTGATGATGACCTCTTTCATGATGATGAGAGAGAGGCTCAGAAACA 567  
Qy 541 AGAGGGAAGATGTTGTTTAAACCAA 567  
Db 568 AGAGGGAAGATGTTGTTTAAACCAA 594

RESULT 15  
AAA44369  
ID AAA44369 standard; cDNA, 895 BP.  
XX  
XX AAA44369;  
XX  
XX 21-AUG-2000 (first entry)  
XX  
XX Human secreted expressed sequence tag SEQ ID NO:944.  
XX  
XX Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;  
XX expressed sequence tag; EST; probe; chemotactic; proliferative;  
XX immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;  
XX thrombolytic; anti-inflammatory; cytostatic; antibacterial; antifungal;  
XX antiviral; antidiabetic; antiaslathmic; vulnery; antiparkinsonian;  
XX antitumor; osteopathic; neuroprotective; nootropic; antipsoriatic;  
XX cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
XX vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
XX insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
XX lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
XX central nervous system disorder; Alzheimer's disease; stroke;  
XX Parkinson's disease; Huntington's disease; coagulation disorder;  
XX haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
XX tumour; infection; depression; psoriasis; ss.  
XX  
XX Homo sapiens.  
XX  
XX MO200021991-A1.  
XX  
XX 20-Apr-2000.  
XX  
XX 15-Oct-1999; 99WO-US24206.  
XX  
XX 15-Oct-1998; 98US-0104436.  
XX  
XX (GEMV ) GENETICS INST INC.  
XX  
XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
XX Merberg D, Treacy M, Bowman MR;  
XX  
XX WPI; 2000-317938/27.  
XX  
XX Isolated polynucleotides, and encoded proteins, comprising secreted  
XX expressed sequence tags (SESTs), useful for treating various disorders  
XX such as autoimmune, infectious, and central nervous system disorders -

XX Claim 1; Page 446; 803bp; English.

CC AAA45926 to AAA45925 represent specifically claimed secreted expressed  
CC sequence tags (ESTs), isolated from human, mouse, chicken and rat  
CC tissue sources. The ESTs can have a range of activities depending on  
CC the tissues they were isolated from. The activities include:  
CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;  
CC cytosratic; antibacterial; antifungal; antiviral; antidiabetic;  
CC antiaesthetic; vulnerary; antitumor; osteopathic; neuroprotective;  
CC neurotropic; antiparkinsonian; antipsoriatic; cerebroprotective;  
CC anticonvulsant; and antidepressant. The ESTs can be used for gene  
CC therapy and in vaccines. The ESTs are useful as probes for the  
CC identification and isolation of full-length cDNAs and genomic DNA  
CC molecules which correspond to the ESTs. Proteins encoded by the ESTs  
CC are useful in assays for determining biological activity and raising  
CC antibodies. They may be useful for treatment of autoimmune disorders  
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
CC osteoporosis, osteoarthritis, central nervous system disorders  
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
CC disease), tumours, bacterial, fungal or viral infections, depression and  
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
CC in the exemplification of the present invention.

XX  
SQ Sequence 895 BP; 274 A; 195 C; 224 G; 201 T; 1 other;

Query Match 96.4%; Score 546.4; DB 21; Length 895;  
Best Local Similarity 99.8%; Pred. No. 5.2e-155;  
Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 20 AAGAAACCACTGGAAAAAATGAATTCCTTATCTTGGCAATTTTCGGTGTTTC 79  
DB 21 AAGAAACCACTGGAAAAAATGAATTCCTTATCTTGGCAATTTTCGGTGTTTC 80  
QY 80 ACCCTTTATCCCTGCTCTGGAAAGCTATATGCAAGATGGCATCTCTAAGAGGACTT 139  
DB 81 ACCCTTTATCCCTGCTCTGGAAAGCTATATGCAAGATGGCATCTCTAAGAGGACTT 140  
QY 140 TTGAAGAAATPAAAGAAATAGCCAGCTGTGAGATGTTGCTAAGCAATCATCAACC 199  
DB 141 TTGAAGAAATPAAAGAAATAGCCAGCTGTGAGATGTTGCTAAGCAATCATCAACC 200  
QY 200 TAGCTGTTTATGTTAAAGCCCAAGACAGATCTTATGAGCGATTGGCACTTCTGGTTGATA 259  
DB 201 TAGCTGTTTATGTTAAAGCCCAAGACAGATCTTATGAGCGATTGGCACTTCTGGTTGATA 260  
QY 260 CTGTTGGACCCGAGCTGAGTGGCTCCAAAGACCTAAGAAAAAGCCATCCAAATTATGTAC 319  
DB 261 CTGTTGGACCCGAGCTGAGTGGCTCCAAAGACCTAAGAAAAAGCCATCCAAATTATGTAC 320  
QY 320 AAAAAGCTGACAGATGGCTGAGAAAGTTCACTTGAGCCAGTGAATACCCCACT 379  
DB 321 AAAAAGCTGACAGATGGCTGAGAAAGTTCACTTGAGCCAGTGAATACCCCACT 380  
QY 380 GGGAGAGGGGAGAAAGATCAGCTGTGATGCTGAGAGCCAAAGATTCAATAGATACCATCC 439  
DB 381 GGGAGAGGGGAGAAAGATCAGCTGTGATGCTGAGAGCCAAAGATTCAATAGATACCATCC 440  
QY 440 TGGGCTTTGGAGAGAGCATTTGGGACTCCTCAAGAGGATTAAGAGGATTTCTGGTGG 499  
DB 441 TGGGCTTTGGAGAGAGCATTTGGGACTCCTCAAGAGGATTAAGAGGATTTCTGGTGG 500  
QY 500 TGACCTCTTTGATGAATCTGAGAGAGGGCTCAGAGCAAGAGGAGATTTGTTT 559  
DB 501 TGACCTCTTTGATGAATCTGAGAGAGGGCTCAGAGCAAGAGGAGATTTGTTT 560  
QY 560 ATAACCAA 567  
DB 561 ATAACCAA 568

Search completed: December 22, 2003, 15:10:10  
Job time : 140.658 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:22:34 ; Search time 1109.8 Seconds  
(without alignments)  
12417.225 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_57\_623

Perfect score: 567  
Sequence: 1 CCGATCAGATTATCTTACA.....AGATTGTTTATATACCA 567

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estm:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rdg:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match length	ID	Description
1	567	100.0	873 14	CB984963 AGENCOURT
2	567	100.0	1038 13	BX355940 BX355940
3	567	100.0	1050 13	BX439467 BX439467
4	567	100.0	1126 13	BX360507 BX360507

5	566.2	99.9	919	13	BX335995	BX335995
6	565.4	99.7	1012	13	BX416896	BX416896
7	563.8	99.4	879	13	BQ218838	BQ218838
8	560	98.8	1024	11	BC012019	BC012019
9	556	98.1	806	12	BI764052	BI764052
10	556	98.1	1081	13	BQ072892	BQ072892
11	555	97.9	745	12	BI462584	BI462584
12	554.4	97.8	783	12	BI754468	BI754468
13	554.4	97.8	909	13	BQ878966	BQ878966
14	553.4	97.6	822	12	BI911772	BI911772
15	553.4	97.6	922	12	BG761741	BG761741
16	553.2	97.6	729	12	BI460469	BI460469
17	553.2	97.6	998	13	BX460463	BX460463
18	552.8	97.5	786	10	BG430966	BG430966
19	552.4	97.4	945	12	BI909780	BI909780
20	551.8	97.3	801	12	BI754143	BI754143
21	547	96.5	836	12	CB958603	CB958603
22	544.8	96.1	1201	13	BX379674	BX379674
23	542.6	95.7	722	12	BI766288	BI766288
24	539.2	95.1	944	12	BI819273	BI819273
25	533	94.0	701	12	BI771109	BI771109
26	528.8	93.3	727	12	BG765433	BG765433
27	528	93.1	756	12	BI757436	BI757436
28	526.6	92.9	862	12	BG772207	BG772207
29	524	92.4	849	10	BG166321	BG166321
30	514	90.7	539	14	CB267661	CB267661
31	514	90.7	562	9	AU279837	AU279837
32	506	89.2	889	10	BE906771	BE906771
33	502.8	88.7	945	13	BQ421059	BQ421059
34	480	84.7	669	12	BI910310	BI910310
35	473.6	83.5	736	12	BI769575	BI769575
36	460	81.1	545	14	CA397585	CA397585
37	456.6	80.5	494	12	BI004612	BI004612
38	456	80.4	751	14	CD58931	CD58931
39	452.6	79.8	717	14	CD000772	CD000772
40	450	79.4	866	13	BU146905	BU146905
41	445	78.5	582	10	BG402129	BG402129
42	430.4	75.9	562	14	CB555139	CB555139
43	430.4	75.9	813	14	CD519131	CD519131
44	417	73.5	498	28	AQ129106	AQ129106
45	416	73.4	653	10	BE218907	BE218907

#### ALIGNMENTS

RESULT 1  
CB984963  
LOCUS  
DEFINITION  
AGENCOURT 13567650 NIH\_MGC\_184 Homo sapiens cDNA clone  
IMAGE:30326745 5', mRNA sequence.  
ACCESSION  
CB984963  
VERSION  
CB984963.1 GI:30279487  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 873)  
NIH-MGC <http://imgc.ncbi.nih.gov/>.  
Natlional Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LMU at:  
<http://image.llnl.gov>  
Plate: NDCM134 row: k column: 10  
High quality sequence stop: 515.





```

Db      558 ATTCATTAAGATAGCAATCTGGGCTTGTGGAGAGCAATGGGACTCTCCAGAGGCAATT 617
Qy      481 ACAGCAGAAAGTTCTGTGTGACCTCTTTCATGAATGCAGAGAGGGCTCAGAAACA 540
Db      618 ACAGCAGAAAGTTCTGTGTGACCTCTTTCATGAATGCAGAGAGGGCTCAGAAACA 677
Qy      541 AGAGGGAAGATTGTTGTTTATTAACCA 567
Db      678 AGAGGGAAGATTGTTGTTTATTAACCA 704

RESULT 3
BX439467 1050 bp mRNA linear EST 15-MAY-2003
LOCUS BX439467 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE010YJ06
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION BX439467 GI:30777746
VERSION BX439467.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1050)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DE010DE03QPI&cluster=4663.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DE010DE03QPI.
FEATURES
Source
1..1050
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE010YJ06"
/tissue_type="PLACENTA"
/clone_id="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 307 a 233 c 271 g 239 t
ORIGIN
Query Match 100.0%; Score 567; DB 13; Length 1050;
Best Local Similarity 100.0%; Pred. No. 4,4e-121;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTTATCAGATTATCTTAACAAGAAACCACTGGAAGAAAAAATGAATTCCTATCTTC 60
Db 100 CTTATCAGATTATCTTAACAAGAAACCACTGGAAGAAAAAATGAATTCCTATCTTC 159
Qy 61 GCATTTTTCGTGTGTTCACCTTTTTCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
Db 160 GCATTTTTCGTGTGTTCACCTTTTTCCTGTGCTCTGGGAAAGCTATATGCAAGAT 219
Qy 121 GGCATCTCTAAGAGAGCTTTTGAAGAATAAAGAAATAGCAGCTGTGAGATGT 180
Db 220 GGCATCTCTAAGAGAGCTTTTGAAGAATAAAGAAATAGCAGCTGTGAGATGT 279
Qy 181 GCTAAGCAATCATCAACTAGCTGTTATGTAAAGCCAGAAAGAGTCTATGAGCA 240

```

```

Db      280 GCTAAGCAATCATCAACTAGCTGTTATGTAAAGCCAGAAAGAGTCTATGAGCA 339
Qy      241 TTGGCACTTGTGTGTACTGTGTGACCGAGCTGAGGGCTCCAGAACTTAAAGAAA 300
Db      340 TTGGCACTTGTGTGTACTGTGTGACCGAGCTGAGGGCTCCAGAACTTAAAGAAA 399
Qy      301 GCCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGGCTGGAAGATTCACTTGAG 360
Db      400 GCCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGGCTGGAAGATTCACTTGAG 459
Qy      361 CCAGTGGAATTAACCCACTGGAGAGGGGAGAAAGATCACTGTGATCTGGAGCCAGA 420
Db      460 CCAGTGGAATTAACCCACTGGAGAGGGGAGAAAGATCACTGTGATCTGGAGCCAGA 519
Qy      421 ATTCATTAAGATAGCAATCTGGGCTTGTGGAGAGCAATGGGACTCTCCAGAGGCAATT 480
Db      520 ATTCATTAAGATAGCAATCTGGGCTTGTGGAGAGCAATGGGACTCTCCAGAGGCAATT 579
Qy      481 ACAGCAGAAAGTTCTGTGTGACCTCTTTCATGAATGCAGAGAGGGCTCAGAAACA 540
Db      580 ACAGCAGAAAGTTCTGTGTGACCTCTTTCATGAATGCAGAGAGGGCTCAGAAACA 639
Qy      541 AGAGGGAAGATTGTTGTTTATTAACCA 567
Db      640 AGAGGGAAGATTGTTGTTTATTAACCA 666

RESULT 4
BX360507 1126 bp mRNA linear EST 05-MAY-2003
LOCUS BX360507 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1071YA19 5-PRIME, mRNA sequence.
ACCESSION BX360507
VERSION BX360507.1 GI:30374439
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1126)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1071YA19QPI&cluster=4663.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0D1071YA19QPI.
FEATURES
Source
1..1126
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1071YA19"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 311 a 249 c 298 g 246 t 22 others
ORIGIN
Query Match 100.0%; Score 567; DB 13; Length 1126;
Best Local Similarity 100.0%; Pred. No. 4,4e-121;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 CCTATCAGATTATCTTAAAGAAAACCACTGAGAAAAAATGAATCTTATCTTC 60
    |||
Db 168 CTTATCAGATTATCTTAAAGAAAACCACTGAGAAAAAATGAATCTTATCTTC 227
Oy 61 GCATTTTTCGGTGTGTTCACCTTTTATCCCTGTCTCTGAGAAAAGCTATATGCAAGAT 120
    |||
Db 228 GCATTTTTCGGTGTGTTCACCTTTTATCCCTGTCTCTGAGAAAAGCTATATGCAAGAT 287
Oy 121 GGCATCTTAAGAGACCTTTTGAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAG 180
    |||
Db 288 GGCATCTTAAGAGACCTTTTGAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAG 347
Oy 181 GCTAAAGCAATCATCACTAGCTGTATGTTAAAGCCAGAAAGATCTTATGAGCGA 240
    |||
Db 348 GCTAAAGCAATCATCACTAGCTGTATGTTAAAGCCAGAAAGATCTTATGAGCGA 407
Oy 241 TTGGCATTCTGTGTGTATCTGTGGACCAAGCTGAGTGGCTCCAGAACCTTGAAGAA 300
    |||
Db 408 TTGGCATTCTGTGTGTATCTGTGGACCAAGCTGAGTGGCTCCAGAACCTTGAAGAA 467
Oy 301 GCCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTCACTGGAG 360
    |||
Db 468 GCCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTCACTGGAG 527
Oy 361 CCAGTGAAGATATCCCACTGGAGAGAGGAGAGAAATCAGCTGTGATGCTGGAGCCAGA 420
    |||
Db 528 CCAGTGAAGATATCCCACTGGAGAGAGGAGAGAAATCAGCTGTGATGCTGGAGCCAGA 587
Oy 421 ATTCTAAGATAGGCAATCTGTGGCTTTGGCAGCAGCATTTGGGACTCTCTCAAGAGCAT 480
    |||
Db 588 ATTCTAAGATAGGCAATCTGTGGCTTTGGCAGCAGCATTTGGGACTCTCTCAAGAGCAT 647
Oy 481 ACAGCAAGAAATTTCTGTGTGTACCTCTTCCATGAATCTGCAGAAAGGCTCTCAGAAACA 540
    |||
Db 648 ACAGCAAGAAATTTCTGTGTGTACCTCTTCCATGAATCTGCAGAAAGGCTCTCAGAAACA 707
Oy 541 AGAGGGAAGATTGTGTTTATTAACCA 567
    |||
Db 708 AGAGGGAAGATTGTGTTTATTAACCA 734

```

```

RESULT 5          919 bp  mRNA  linear  EST 02-MAY-2003
BX335995          clone CS0D1022YJ16 5-PRIME, mRNA sequence.
LOCUS              BX335995
ACCESSION          BX335995
VERSION            BX335995.1 GI:30339459
KEYWORDS            EST.
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE           1 (bases 1 to 919)
AUTHORS             Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE               Full-length cDNA libraries and normalization
JOURNAL             Unpublished
COMMENT             Contact: Genoscope
                    Genoscope - Centre National de Sequencage
                    BP 191 91006 Evry cedex - France
                    Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
                    Library was constructed by Life Technologies, a division of
                    Invitrogen. This sequence belongs to sequence cluster 4663.f For
                    more information about this cluster, see
                    http://www.genoscope.cns.fr/
                    cgl-bi/cluster.cgi?seq=CS0D1022DE080P1c1cluster=4663.f. Contact :
                    Feng Liang Email : fliang@lifetech.com URL :
                    http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
                    Faraday Avenue Genoscope sequence ID : CS0D1022DE080P1.

```

FEATURES  
source

1. 919  
/organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1022YJ16"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      273 a      206 c      226 g      208 t      6 others
ORIGIN

```

```

Query Match      99.9%; Score 566.2; DB 13; Length 919;
Best Local Similarity 99.6%; Pred. No. 6.8e-121;
Matches 565; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 CCTATCAGATTATCTTAAAGAAAACCACTGAGAAAAAATGAATCTTATCTTC 60
    |||
Db 84 CTTATCAGATTATCTTAAAGAAAACCACTGAGAAAAAATGAATCTTATCTTC 143
Oy 61 GCATTTTTCGGTGTGTTCACCTTTTATCCCTGTCTCTGAGAAAAGCTATATGCAAGAT 120
    |||
Db 144 GCATTTTTCGGTGTGTTCACCTTTTATCCCTGTCTCTGAGAAAAGCTATATGCAAGAT 203
Oy 121 GGCATCTTAAGAGACCTTTTGAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAG 180
    |||
Db 204 GGCATCTTAAGAGACCTTTTGAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAG 263
Oy 181 GCTAAAGCAATCATCACTAGCTGTATGTTAAAGCCAGAAAGATCTTATGAGCGA 240
    |||
Db 264 GCTAAAGCAATCATCACTAGCTGTATGTTAAAGCCAGAAAGATCTTATGAGCGA 323
Oy 241 TTGGCATTCTGTGTGTATCTGTGGACCAAGCTGAGTGGCTCCAGAACCTTGAAGAA 300
    |||
Db 324 TTGGCATTCTGTGTGTATCTGTGGACCAAGCTGAGTGGCTCCAGAACCTTGAAGAA 383
Oy 301 GCCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTCACTGGAG 360
    |||
Db 384 GCCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGCTGAGAAAGTTCACTGGAG 443
Oy 361 CCAGTGAAGATATCCCACTGGAGAGAGGAGAGAAATCAGCTGTGATGCTGGAGCCAGA 420
    |||
Db 444 CCAGTGAAGATATCCCACTGGAGAGAGGAGAGAAATCAGCTGTGATGCTGGAGCCAGA 503
Oy 421 ATTCTAAGATAGGCAATCTGTGGCTTTGGCAGCAGCATTTGGGACTCTCTCAAGAGCAT 480
    |||
Db 504 ATTCTAAGATAGGCAATCTGTGGCTTTGGCAGCAGCATTTGGGACTCTCTCAAGAGCAT 563
Oy 481 ACAGCAAGAAATTTCTGTGTGTACCTCTTCCATGAATCTGCAGAAAGGCTCTCAGAAACA 540
    |||
Db 564 ACAGCAAGAAATTTCTGTGTGTACCTCTTCCATGAATCTGCAGAAAGGCTCTCAGAAACA 623
Oy 541 AGAGGGAAGATTGTGTTTATTAACCA 567
    |||
Db 624 AGAGGGAAGATTGTGTTTATTAACCA 650

```

```

RESULT 6          1012 bp  mRNA  linear  EST 13-MAY-2003
BX416896          clone CS0D001YF04
LOCUS              BX416896
ACCESSION          BX416896
VERSION            BX416896.1 GI:30650313
KEYWORDS            EST.
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE           1 (bases 1 to 1012)
AUTHORS             Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE               Full-length cDNA libraries and normalization
JOURNAL             Unpublished

```

## COMMENT

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4663.f. For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DE001DC02QPLcluster=4663.f. Contact :  
Peng Liang Email : lliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS0DE001DC02QPL.

## FEATURES

## Source

1. 1012  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE001F04"  
/issue\_type="PLACENTA"  
/clone\_lib="Homo sapiens PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."  
Library was not normalized."  
BASE COUNT 281 a 240 c 275 g 215 t 1 others  
ORIGIN

Query Match 99.7%; Score 565.4; DB 13; Length 1012;  
Best Local Similarity 99.8%; Pred. No. 1e-120;  
Matches 566; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CCTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTATCTTC 60  
213 CCTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTATCTTC 272  
61 GCATTTTTCGGTGTGTACCTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120  
273 GCATTTTTCAGTGTGTACCTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 332  
121 GGCATCTCTAAGAGACTTTTGAAGAATAAAGAAATAGCCAGCTGTGAGATTT 180  
333 GGCATCTCTAAGAGACTTTTGAAGAATAAAGAAATAGCCAGCTGTGAGATTT 392  
181 GCTAAGCAATCATCAACTGAGCTTTATGTAAAGCCGAGACAGATCTTATGAGCGA 240  
393 GCTAAGCAATCATCAACTGAGCTTTATGTAAAGCCGAGACAGATCTTATGAGCGA 452  
241 TTGGCACTTCTGTGTAAGTGTGAGCCAGACTGAGTGTCCAGAACTTGAAGAAA 300  
453 TTGGCACTTCTGTGTAAGTGTGAGCCAGACTGAGTGTCCAGAACTTGAAGAAA 512  
301 GCCATCCAAATTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAATTCACCTGAG 360  
513 GCCATCCAAATTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAATTCACCTGAG 572  
361 CCAATGAGAAATACCCCACTGGGAGAGGAGAGAAATAGCTGTATCTGAGCCAGAA 420  
573 CCAATGAGAAATACCCCACTGGGAGAGGAGAGAAATAGCTGTATCTGAGCCAGAA 632  
421 ATTCAATAGATAGCATCTCTGGGCTTTGGAGAGCATTTGGGAACTCTCCAGAAAGGCA 480  
633 ATTCAATAGATAGCATCTCTGGGCTTTGGAGAGCATTTGGGAACTCTCCAGAAAGGCA 692  
481 ACAGCAGAAAGTTGTGTGTGTAAGTGTGAGCAAGATGAGAGAGGCTCTAGAAACA 540  
693 ACAGCAGAAAGTTGTGTGTGTAAGTGTGAGCAAGATGAGAGAGGCTCTAGAAACA 752  
541 AGAGGAGAAATTTGTTTATTAACCA 567  
753 AGAGGAGAAATTTGTTTATTAACCA 779

## RESULT 7

B0218838 879 bp mRNA linear EST 02-MAY-2002  
LOCUS AGENCOURT 7546057 NIH\_MGC\_70 Homo sapiens cDNA clone IMAGE:6025034  
DEFINITION 5', mRNA sequence.  
ACCESSION B0218838  
VERSION B0218838.1 GI:20400238  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

## REFERENCE

NIH-MGC http://mgc.nci.nih.gov/

## AUTHORS

1 (bases 1 to 879)

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgabs@wmail.nih.gov  
Tissue Procurement: ATCC  
DNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LHAM3236 row: 1 column: 03  
High quality sequence stop: 589.

## FEATURES

## Source

1. 879  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6025034"  
/issue\_type="epithelioid carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_70"  
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.1 kb. Library constructed by Life  
Technologies."  
BASE COUNT 239 a 216 c 235 g 189 t  
ORIGIN

Query Match 99.4%; Score 563.8; DB 13; Length 879;  
Best Local Similarity 99.6%; Pred. No. 2.5e-120;  
Matches 565; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 CCTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTATCTTC 60  
103 CCTATCAGATTATCTTAACAAGAAACCACTGGAAAAAATGAATTCCTATCTTC 162  
61 GCATTTTTCGGTGTGTACCTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120  
163 GCATTTTTCGGTGTGTACCTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 222  
121 GGCATCTCTAAGAGACTTTTGAAGAATAAAGAAATAGCCAGCTGTGAGATTT 180  
223 GGCATCTCTAAGAGACTTTTGAAGAATAAAGAAATAGCCAGCTGTGAGATTT 282  
181 GCTAAGCAATCATCAACTGAGCTTTATGTAAAGCCGAGACAGATCTTATGAGCGA 240  
283 GCTAAGCAATCATCAACTGAGCTTTATGTAAAGCCGAGACAGATCTTATGAGCGA 342  
241 TTGGCACTTCTGTGTAAGTGTGAGCCAGACTGAGTGTCCAGAACTTGAAGAAA 300  
343 TTGGCACTTCTGTGTAAGTGTGAGCCAGACTGAGTGTCCAGAACTTGAAGAAA 402  
301 GCCATCCAAATTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAATTCACCTGAG 360  
403 GCCATCCAAATTATGTACCAAAACCTGAGCAAGATGGGCTGAGAAAATTCACCTGAG 462  
361 CCAATGAGAAATACCCCACTGGGAGAGGAGAGAAATAGCTGTATCTGAGCCAGAA 420

Db 463 CCAGTGAAGATACCCCACTGGAGAGAGGAGAGAGATCAGCTGTGATCTGGAGCCAAAGA 522  
Oy 421 ATTCAATAGATAGCCATCTGGGCTCTTGGACAGACATTTGGAGACCTCTCCGAAGGCATT 480  
Db 523 ATTCTAAGATATGACATCTGGGCTCTTGGACAGACATTTGGAGACCTCTCCGAAGGCATT 582  
Oy 481 ACAGCAGAGATTTCTGGTGTGACCTCTTTCATGATGACAGAGAGAGGAGCTCAGAAACA 540  
Db 583 ACAGCAGAGATTTCTGGTGTGACCTCTTTCATGATGACAGAGAGAGGAGCTCAGAAATC 642  
Oy 541 AGAGGAGAGATTTGTTTATTAACCA 567  
Db 643 AGAGGAGAGATTTGTTTATTAACCA 669

RESULT 8  
LOCUS BC012019 1024 bp mRNA linear HTC 04-MAR-2003  
DEFINITION Homo sapiens, similar to plasma glutamate carboxypeptidase, clone  
IMAGE:4455631, mRNA.  
BC012019  
ACCESSION BC012019.1 GI:15080558  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1024)  
Strausberg, R.  
Direct Submission  
Submitted (30-JUN-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.mci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gumarane, P.H., Garcia, A.M., Lu, X., Huliy, S.W., Louisse, H.,  
Kovals, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nantavali,  
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 28 Row: b Column: 11  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 5174626  
This clone has the following problem: retained intron.  
Location/Qualifiers  
1..1024  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4455631"  
/tissue\_type="Kidney, hypernephroma"  
/clone\_id="NIH\_MGC\_69"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
BASE COUNT 317 a 210 c 256 g 241 t  
ORIGIN

Query Match 98.8%; Score 560; DB 11; Length 1024;  
Best local Similarity 100.0%; Pred. No. 1.9e-119;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8 GATTATCTTAACAAGAAACCAACTGGAAGAAAAAATGAAATTCCTATCTTCGATTTT 67

Db 1 GATTATCTTAACAAGAAACCAACTGGAAGAAAAAATGAAATTCCTATCTTCGATTTT 60  
Oy 68 TCGGTGGTTCACCTTTTATCCCTGCTCTGGGAAGCATATGCAAGATGGCATCT 127  
Db 61 TCGGTGGTTCACCTTTTATCCCTGCTCTGGGAAGCATATGCAAGATGGCATCT 120  
Oy 128 CTAAGAGGACCTTTGAAGAAATAAAGAAATAGCAGCTGTGAGATGTTCTAAAG 187  
Db 121 CTAAGAGGACCTTTGAAGAAATAAAGAAATAGCAGCTGTGAGATGTTCTAAAG 180  
Oy 188 CAATCATCAACCTAGCTGTTTATGTAAGCCGAAACAATCTATGAGCATGGCAC 247  
Db 181 CAATCATCAACCTAGCTGTTTATGTAAGCCGAAACAATCTATGAGCATGGCAC 240  
Oy 248 TTCTGGTTGATACCTGTGACCCAGACTGAGTGTCTCAAGAACCTGAAAAACCATCC 307  
Db 241 TTCTGGTTGATACCTGTGACCCAGACTGAGTGTCTCAAGAACCTGAAAAACCATCC 300  
Oy 308 AAATTATGTACCAAAACCTGAGCAAGATGAGGCTGAGAAAGTTCACTGAGCCAGTGA 367  
Db 301 AAATTATGTACCAAAACCTGAGCAAGATGAGGCTGAGAAAGTTCACTGAGCCAGTGA 360  
Oy 368 GAATACCCCACTGGAGAGAGGAGAGAAATCAGCTGTGATGCTGAGCCAGAAATTCATA 427  
Db 361 GAATACCCCACTGGAGAGAGGAGAGAAATCAGCTGTGATGCTGAGCCAGAAATTCATA 420  
Oy 428 AGATAGCCATCTCTGGGCTCTTGGACGACGATTTGGGATCTCTCCAGAAAGCATTCACACAG 487  
Db 421 AGATAGCCATCTCTGGGCTCTTGGACGACGATTTGGGATCTCTCCAGAAAGCATTCACACAG 480  
Oy 488 AAGTTCTGTGTGATACCTCTTTCATGAACTGAGAGAGGAGGCTCAGAAAGCAAGAGGA 547  
Db 481 AAGTTCTGTGTGATACCTCTTTCATGAACTGAGAGAGGAGGCTCAGAAAGCAAGAGGA 540  
Oy 548 AGATTTGTTTATTAACCA 567  
Db 541 AGATTTGTTTATTAACCA 560

RESULT 9  
LOCUS B1764052 806 bp mRNA linear EST 25-SEP-2001  
DEFINITION 603043258F1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5183559 5',  
mRNA sequence.  
B1764052  
ACCESSION B1764052.1 GI:15755630  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 806)  
NIH-MGC <http://mgc.mci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: L1AM11458 Row: c Column: 16  
High quality sequence stop: 804.  
Location/Qualifiers  
1..806  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5183559"

/lab\_host="DH10B"  
/clone\_1ib="NIH MGC 118"  
/note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon, 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH MGC Library."

BASE COUNT  
ORIGIN

221 a 194 c 213 g 178 t

Query Match 98.1%; Score 556; DB 12; Length 806;  
Best Local Similarity 99.8%; Pred. No. 1.6e-118;  
Matches 567; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CCTATCAGATTATCTTAAAGAAACCACTGGAAAAAATGAATTCCTATCTTC 60  
DB 115 CCTATCAGATTATCTTAAAGAAACCACTGGAAAAAATGAATTCCTATCTTC 174  
QY 61 GCATTTTTCGGTGGTGTTCACCTTTTATCCCTGTCTCTGGAAAGCTATATGCAAGAT 120  
DB 175 GCATTTTTCGGTGGTGTTCACCTTTTATCCCTGTCTCTGGAAAGCTATATGCAAGAT 234  
QY 121 GCATCTCTTAAGAGAGACTTTTGAAGAAATTAAGAAAGATAGCCAGCTGTGAGATGTT 180  
DB 235 GCATCTCTTAAGAGAGACTTTTGAAGAAATTAAGAAAGATAGCCAGCTGTGAGATGTT 294  
QY 181 GCTAAGGCAATCATCAACCTAGCTGTATGTATGTAAGCCAGAAACAGATCCTATGAGCGA 240  
DB 295 GCTAAGGCAATCATCAACCTAGCTGTATGTATGTAAGCCAGAAACAGATCCTATGAGCGA 354  
QY 241 TTGGCACTTCTGTGTGATCTGTGTGACCACTGAGTGGCTCCAGAACTTGAAGAAA 300  
DB 355 TTGGCACTTCTGTGTGATCTGTGTGACCACTGAGTGGCTCCAGAACTTGAAGAAA 414  
QY 301 GCCATCCAAATTATATGACCAAAACCTGACAGAGATGGGCTGGAGAAAGTTCACTGGAG 360  
DB 415 GCCATCCAAATTATATGACCAAAACCTGACAGAGATGGGCTGGAGAAAGTTCACTGGAG 474  
QY 361 CCAGTGAAGATTAACCCCACTGGAGAGAGGAGAAAGAAATCAGCTGTGATGCTGAGCCAGA 420  
DB 475 CCAGTGAAGATTAACCCCACTGGAGAGAGGAGAAAGAAATCAGCTGTGATGCTGAGCCAGA 534  
QY 421 ATTCAATTAAGATGAGCATCTGTGGTCTTGGAGAGAGATGGGAACTCTCCAGAAAGCATT 480  
DB 535 ATTCAATTAAGATGAGCATCTGTGGTCTTGGAGAGAGATGGGAACTCTCTCCAGAAAGCATT 594  
QY 481 ACAGCAGAAAGTTGTGTGTGACCTCTTTGATGATGATGC-AGAAGAGGGCTTCAGAGC 539  
DB 595 ACAGCAGAAAGTTGTGTGTGACCTCTTTGATGATGATGCAGAGAGAGGGCTTCAGAGC 654  
QY 540 AAGAGGAGAAATGTTGTTTATTAACCA 567  
DB 655 AAGAGGAGAAATGTTGTTTATTAACCA 682

RESULT 10  
LOCUS B0072892 1081 bp mRNA linear EST 02-APR-2002  
DEFINITION AGENCOURT\_6763138 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5756387  
ACCESSION B0072892  
VERSION B0072892.1 GI:19901938  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1081)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LLM412796 row: 0 column: 12  
High quality sequence stop: 658.

## FEATURES

## Source

Location/Qualifiers

1..1081  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/mol\_type="mRNA"  
/clone="IMAGE:5756387"  
/issue\_type="leukocyte"  
/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH MGC Library."

BASE COUNT 319 a 256 c 275 g 225 t 2 others

ORIGIN

Query Match 98.1%; Score 556; DB 13; Length 1081;  
Best Local Similarity 99.8%; Pred. No. 1.6e-118;  
Matches 567; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CCTATCAGATTATCTTAAAGAAACCAACT-GGAAAAAATGAATTCCTATCTT 59  
DB 123 CCTATCAGATTATCTTAAAGAAACCAACTGGAAAAAATGAATTCCTATCTT 182  
QY 60 CGATTTTTCGGTGGTGTTCACCTTTTATCCCTGTCTCTGGAAAGCTATATGCAAGAA 119  
DB 183 CGATTTTTCGGTGGTGTTCACCTTTTATCCCTGTCTCTGGAAAGCTATATGCAAGAA 242  
QY 120 TTGGCATCTTAAGAGAGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGT 179  
DB 243 TTGGCATCTTAAGAGAGACTTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGT 302  
QY 180 TTGCTAAGCAATCATCAACCTAGCTGTATGTGTAAGCCAGAAACAGATCTATGAGCG 239  
DB 303 TTGCTAAGCAATCATCAACCTAGCTGTATGTGTAAGCCAGAAACAGATCTATGAGCG 362  
QY 240 ATTGGCACTTCTGTGTGATCTGTGAGCCCAAGATGAGGCTCCAGAAAGCTTAAGAAA 299  
DB 363 ATTGGCACTTCTGTGTGATCTGTGAGCCCAAGATGAGGCTCCAGAAAGCTTAAGAAA 422  
QY 300 AGCCATCCAAATTATATGACCAAAACCTGACAGAAATGGGCTGGAGAAAGTTCACTGGA 359  
DB 423 AGCCATCCAAATTATATGACCAAAACCTGACAGAAATGGGCTGGAGAAAGTTCACTGGA 482  
QY 360 GCCAGTGAAGTAACCCCACTGGAGAGAGGAGAGAAATCAGCTGTGATGCTGAGCCAG 419  
DB 483 GCCAGTGAAGTAACCCCACTGGAGAGAGGAGAGAAATCAGCTGTGATGCTGAGCCAG 542  
QY 420 AATTCAATTAAGATGAGCATCTGTGGTCTTGGAGACAGATTTGGAGACTCTCCAGAAAGCAT 479  
DB 543 AATTCAATTAAGATGAGCATCTGTGGTCTTGGAGACAGATTTGGAGACTCTCCAGAAAGCAT 602  
QY 480 TACAGCAGAAAGTTCTGGTGTGACCTCTTTCATGTAACCTCAGAGAAAGGCTTCAGAGC 539

Db 603 TACGACGAAGTCTGTGGTGAAGCTCTTTCGATGAGTGAAGGAGGCTCAGAAC 662

Qy 540 AAGAGGGAAGATTGTTTATTAACCA 567

Db 663 AAGGAGGAAGATTGTTTATTAACCA 690

RESULT 11

LOCUS B1462584 745 bp mRNA linear EST 21-AUG-2001

DEFINITION 60302050F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5268937 5', mRNA sequence.

ACCESSION B1462584

VERSION B1462584.1 GI:15253240

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 745) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC)

AUTHORS Unpublished

TITLE Contact: Robert Strausberg, Ph.D.

JOURNAL Email: cgabs-remail.nih.gov

COMMENT Tissue Procurement: Mikhail Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: L14M1678 row: 1 column: 02 High quality sequence stop: 741. Location/Qualifiers

FEATURES

source 1. 745

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5268937"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_97"

/note="Organ: testis; Vector: pBluescript (modified pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (Gcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTT-3', size selected for average insert size 2.2 kb and normalized to 10^5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 213 a 169 c 198 g 164 t 1 others

ORIGIN

Query Match 97.9%; Score 555; DB 12; Length 745;

Best Local Similarity 99.8%; Pred. No. 2.7e-118;

Matches 565; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 CCGATCAGATTATCTTAACAAGAAACCAACTGGAAAAAATGAATTCCTATCTC 60

Db 87 CCTATCAGATTATCTTAACAAGAAACCAACTGGAAAAAATGAATTCCTATCTC 146

Qy 61 GCATTTTCGGTGTGTCACTTTATCCCTGTCTCTGGAAAGCTATATGCAAGAT 120

Db 147 GCATTTTCGGTGTGTCACTTTATCCCTGTCTCTGGAAAGCTATATGCAAGAT 206

Qy 121 GGCATCTCTAAGAGACTTTTGAAGAATAAAGAAATAGCCAGCTGTGAGATGT 180

Db 207 GGCATCTCTAAGAGACTTTTGAAGAATAAAGAAATAGCCAGCTGTGAGATGT 266

Qy 181 GCTAAGCAATCATCACTAGCTGTTTATAGTAAAGCCCAAGACATCTTATAGCA 240

Db 267 GCTAAGCAATCATCACTAGCTGTTTATAGTAAAGCCCAAGACATCTTATAGCA 326

Qy 241 TTGGCACTCTGTTATATCTGTGACCCAGACTAGTGTCTCAAGAACTAGAAAA 300

Db 327 TTGGCACTCTGTTATATCTGTGACCCAGACTAGTGTCTCAAGAACTAGAAAA 386

Qy 301 GCCATCAATTAATATGACAAACCTGCAGCAAGATGGGCTGGAGAAAGTTCACCTGAG 360

Db 387 GCCATCAATTAATATGACAAACCTGCAGCAAGATGGGCTGGAGAAAGTTCACCTGAG 446

Qy 361 CCAAGTGAATATCCCACTGGAGAGAGGAGAAAGAAATCACTGTGATGCTGAGCCAGA 420

Db 447 CCAAGTGAATATCCCACTGGAGAGAGGAGAAAGAAATCACTGTGATGCTGAGCCAGA 506

Qy 421 ATTCATTAAGTATGACCATCTGCTGTGGCAGAGCATTTGGACCTCTCAAGAGCAT 480

Db 507 ATTCATTAAGTATGACCATCTGCTGTGGCAGAGCATTTGGACCTCTCAAGAGCAT 566

Qy 481 ACAGCAAGATTCTGTGGTGAACCTCTTTCGATGAATGACAGAGAGGCTCAGAGCA 540

Db 567 ACAGCAAGATTCTGTGGTGAACCTCTTTCGATGAATGACAGAGAGGCTCAGAGCA 626

Qy 541 AGAGGGAAGATTGTTTATTAACCA 567

Db 627 AGAGGGAAGATTGTTTATTAACCA 652

RESULT 12

LOCUS B1754468 783 bp mRNA linear EST 25-SEP-2001

DEFINITION 603022607F1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5193385 5', mRNA sequence.

ACCESSION B1754468

VERSION B1754468.1 GI:15746046

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 783) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC)

AUTHORS Unpublished

TITLE Contact: Robert Strausberg, Ph.D.

JOURNAL Email: cgabs-remail.nih.gov

COMMENT Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: L14M1483 row: m column: 02 High quality sequence stop: 783. Location/Qualifiers

FEATURES

source 1. 783

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5193385"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_114"

/note="Organ: Brain; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH\_MGC Library."

BASE COUNT 223 a 188 c 197 g 175 t

ORIGIN

Query Match 97.8%; Score 554.4; DB 12; Length 783;  
Best Local Similarity 99.6%; Pred. No. 3.7e-118;  
Matches 566; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy	1	CCATACAGATTATCTTAAACAAGAAAAACCACTGAAAAAATGAAATTCCTTATCTTC	60
Db	73	CCATACAGATTATCTTAAACAAGAAAAACCACTGAAAAAATGAAATTCCTTATCTTC	132
Qy	61	GCATTTTTCGGTGGTTCACCTTTTATCCCTGCTCTCGGAAAGCTATATGCAAGAT	120
Db	133	GCATTTTTCGGTGGTTCACCTTTTATCCCTGCTCTCGGAAAGCTATATGCAAGAT	192
Qy	121	GGCATCTCTAAGAGACCTTTGAGAAATTAAGAAATAGCCAGCTGTGAGATGTT	180
Db	193	GGCATCTCTAAGAGACCTTTGAGAAATTAAGAAATAGCCAGCTGTGAGATGTT	252
Qy	181	GCTAAAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCAGAACATCTATGAGCA	240
Db	253	GCTAAAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCAGAACATCTATGAGCA	312
Qy	241	TTGGCATCTTGGTGGTGTGATCTGTGAGACCACTGATGCTGCTCAAGACCTAGAAA	300
Db	313	TTGGCATCTTGGTGGTGTGATCTGTGAGACCACTGATGCTGCTCAAGACCTAGAAA	372
Qy	301	GCATCCCAATTTATGTATCAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGAG	360
Db	373	GCATCCCAATTTATGTATCAAAACCTGAGCAAGATGGGCTGAGAAAGTTCACTGAG	432
Qy	361	CCAGTGAAGATATACCCCACTGGAGAGAGGAGAGAAATCACTGTATGCTGAGCAAGA	420
Db	433	CCAGTGAAGATATACCCCACTGGAGAGAGGAGAGAAATCACTGTATGCTGAGCAAGA	492
Qy	421	ATTGATTAAGATACCATCTCTGGCTTTGAGCAGCAGCATTTGGAGCTCTCCAGAAAGCAT	480
Db	493	ATTGATTAAGATACCATCTCTGGCTTTGAGCAGCAGCATTTGGAGCTCTCCAGAAAGCAT	552
Qy	481	ACAGAGAAAGTTGGTGGTGTGACCTCTTGATGTAAGTGAAGAGAGGCTCCAGAAACA	540
Db	553	ACAGAGAAAGTTGGTGGTGTGACCTCTTGATGTAAGTGAAGAGAGGCTCCAGAAACA	612
Qy	541	AGAGGGAAG-ATTGTTGTTTATTAACCA 567	
Db	613	AGAGGGAAGCATGTGCTTTATTAACCA 640	

RESULT 13  
B0878966 909 bp mRNA linear EST 16-AUG-2002  
LOCUS B0878966  
DEFINITION AGNCOURT 8183107 lupski\_dorsal\_root\_ganglion Homo sapiens cDNA  
clone IMAGE:6184129 5', mRNA sequence.  
ACCESSION B0878966  
VERSION B0878966.1 GI:22270974  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.  
TITLE NIH-MGC http://mgi.nci.nih.gov/.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: Dr. James R. Lupski  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Cloning distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov

Plate: LLM13572 row: n column: 02

FEATURES

High quality sequence stop: 696.  
Location/Qualifiers  
1..909  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6184129"  
/sex="male"  
/tissue\_type="dorsal root ganglia"  
/dev\_stage="adult, 36 yr"  
/lab\_host="DH10B"  
/clone\_id="lupski\_dorsal\_root\_ganglion"  
/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:  
NotI; Site 2: SalI; cDNA made by oligo-dT priming.  
Directionally cloned using the following adaptors:  
5'-TCGACCCAGGCGCTCG-3' and  
5'-GACTAGTCTTGAATGCGAGCGGCCCTT(15)-3'. Size selected >  
1 kb for average insert length 1.7 kb. This is a primary  
library, non-amplified. Library constructed by Life  
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
College of Medicine) and is available through Life  
Technologies."

Qy	1	CCATACAGATTATCTTAAACAAGAAAAACCACT-GGAAAAAATGAAATTCCTTATCTT	59
Db	21	CCATACAGATTATCTTAAACAAGAAAAACCACTGGAAAAAATGAAATTCCTTATCTT	80
Qy	60	GCATTTTTCGGTGGTTCACCTTTTATCCCTGCTCTGGAAGCTATATGCAAGAA	119
Db	81	GCATTTTTCGGTGGTTCACCTTTTATCCCTGCTCTGGAAGCTATATGCAAGAA	140
Qy	120	TGGCATCTTGAAGAGACCTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGT	179
Db	141	TGGCATCTTGAAGAGACCTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGT	200
Qy	180	TGCTTAAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCAGAACATCTATGAGCG	239
Db	201	TGCTTAAGCAATCATCAACCTAGCTGTTTATGTTAAAGCCAGAACATCTATGAGCG	260
Qy	240	ATTGGCACTTCTGTTGATCTGTTGAGCCAGACTGATGCTCCAGAAAGCTAGAAA	299
Db	261	ATTGGCACTTCTGTTGATCTGTTGAGCCAGACTGATGCTCCAGAAAGCTAGAAA	320
Qy	300	AGCCATCCAAATTTATGTAACCAAAACCTGACAGCAAGATGGGTGAGAAAGTTCACTGGA	359
Db	321	AGCCATCCAAATTTATGTAACCAAAACCTGACAGCAAGATGGGTGAGAAAGTTCACTGGA	380
Qy	360	GCCAGTGAAGATATACCCCACTGGAGAGAGGAGAGAAATCACTGTGATGCTGAGCCAG	419
Db	381	GCCAGTGAAGATATACCCCACTGGAGAGAGGAGAGAAATCACTGTGATGCTGAGCCAG	440
Qy	420	AATTCATTAAGATACCATCTCTGGCTTTGAGCAGCAGCATTTGGAGCTCTCCAGAAAGCAT	479
Db	441	AATTCATTAAGATACCATCTCTGGCTTTGAGCAGCAGCATTTGGAGCTCTCCAGAAAGCAT	500
Qy	480	TACAGCAAGATTTCTGGTGGTGTGACCTCTTTCAGAGAACTGAGAGAAAGGCTCAGAGC	539
Db	501	TACAGCAAGATTTCTGGTGGTGTGACCTCTTTCAGAGAACTGAGAGAAAGGCTCAGAGC	560
Qy	540	AAGAGGGAAGATTTGTTTATTAACCA 567	
Db	561	AAGAGGGAAGATTTGTTTATTAACCA 588	

RESULT 14  
B1911772 822 bp mRNA linear EST 16-OCT-2001  
LOCUS B1911772



```

DEFINITION 603065377F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5214451 5',
            mRNA sequence.
ACCESSION  B1911772
VERSION    B1911772.1 GI:16175644
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 822)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            JOURNAL Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLN at:
            http://image.llnl.gov
            Plate: LLM11538 row: j column: 20
            High quality sequence start: 31
            High quality sequence stop: 795.
            Location/Qualifiers
                1..822
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5214451"
                /tissue_type="leukocyte"
                /lab_host="DH10B"
                /clone_id="NIH_MGC_118"
                /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
                (destroyed); RNA source leukocytes from anonymous pool of
                non-activated adult donors. Library is oligo-dT primed
                and directionally cloned (EcoRV site is destroyed upon
                cloning). Average insert size 1.7 kb. Insert size range
                1.2-3.3 kb. Library is normalized and enriched for
                full-length clones and was constructed by C. Gruber
                (Invitrogen). Research Genetics tracking code 027. Note:
                this is a NIH MGC Library."
BASE COUNT  232 a 192 c 215 g 183 t
ORIGIN
Query Match 97.6%; Score 553.4; DB 12; Length 822;
Best Local Similarity 99.6%; Pred. No. 6.4e-118;
Matches 565; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CCTATCAGATTATCTTAACAAGAAACCACTGGAGAAAAAATGAATTCCTTATCTTC 60
DB 95 CTTATCAGATTATCTTAACAAGAAACCACTGGAGAAAAAATGAATTCCTTATCTTC 154
QY 61 GCATTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
DB 155 GCATTTTCGGTGTGTTACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 214
QY 121 GGCATCTTAAGAGACCTTTTGAAGAATAAAGAAATAGCCAGCTGTGGAGATTT 180
DB 215 GGCATCTCTTAAGAGACCTTTTGAAGAATAAAGAAATAGCCAGCTGTGGAGATTT 274
QY 181 GCTAAGCAATCATCACTAGCTGTTATAGTTAAAGCCAGAACAGATCCATAGAGCA 240
DB 275 GCTAAGCAATCATCACTAGCTGTTATAGTTAAAGCCAGAACAGATCCATAGAGCA 334
QY 241 TTGGCACTTCTGTTGATCTGTGGACCCAGACTGATGTGCTCCAGAACCTAGAAAAA 300
DB 335 TTGGCACTTCTGTTGATCTGTGGACCCAGACTGATGTGCTCCAGAACCTAGAAAAA 394
QY 301 GGCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGCGCTGGAGAAATTCCTGGAG 360
DB 395 GGCATCCAAATTATGTACCAAAACCTGCAGCAAGATGGCGCTGGAGAAATTCCTGGAG 454

```

```

QY 361 CCACTGAGAAATACCCCACTGGAGAGAGGAGAAAGATCAGCTGTATGTGAGCCAGA 420
DB 455 CCACTGAGAAATACCCCACTGGAGAGAGGAGAAAGATCAGCTGTATGTGAGCCAGA 514
QY 421 ATTCATTAAGATTAAGCATCTGGGTCTTTGGAGCAGCATTTGGAGCTCTCCAGAGGAT 480
DB 515 ATTCATTAAGATTAAGCATCTGGGTCTTTGGAGCAGCATTTGGAGCTCTCCAGAGGAT 574
QY 481 ACAGCAAGATTCGTGGTGTGACCTCTTTGATTAATTCAGAGAAAGGCGCTCAGAGCA 540
DB 575 ACAGCAAGATTCGTGGTGTGACCTCTTTGATTAATTCAGAGAAAGGCGCTCAGAGCA 634
QY 541 AGAGCAAGATTCGTGGTGTGATTAACCAA 567
DB 635 AGAGCAAGATTCG-TGTTTAAACCAA 660

RESULT 15
BG761741
LOCUS      602717936F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4841505 5',
DEFINITION mRNA sequence.
ACCESSION  BG761741
VERSION    BG761741.1 GI:14072394
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 922)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            JOURNAL Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: ATCC/DCTD/DRP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLN at:
            http://image.llnl.gov
            Plate: LNCM1674 row: g column: 10
            High quality sequence stop: 796.
            Location/Qualifiers
                1..922
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4841505"
                /tissue_type="melanotic melanoma, high MDR (cell line)"
                /lab_host="DH10B (phage-resistant)"
                /clone_id="NIH_MGC_49"
                /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                into EcoRV/XhoI sites using the following 5' adaptor:
                GGCACAGC(G). Size-selected >500bp for average insert size
                1.8kb. Library constructed by Ling Hong in the laboratory
                of Gerald M. Rubin (University of California, Berkeley)
                using ZAP-cDNA synthesis kit (Stratagene) and Superscript
                II RT (Life Technologies). Note: this is a NIH_MGC
                Library."
BASE COUNT  279 a 204 c 227 g 211 t 1 others
ORIGIN
Query Match 97.6%; Score 553.4; DB 12; Length 922;
Best Local Similarity 99.6%; Pred. No. 6.4e-118;
Matches 565; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTTATCAGATTATCTTAACAAGAAACCACTGGAGAAAAAATGAATTCCTTATCTTC 60
DB 22 CTTATCAGATTATCTTAACAAGAAACCACTGGAGAAAAAATGAATTCCTTATCTTC 81

```



```
QY 61 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGAAAGCTATATGCAGAAAT 120
    |||||
Db 82 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGAAAGCTATATGCAGAAAT 141
    |||||
QY 121 GGCATCTCTTAAGAGAGACTTTTGAAGAAATAAAGAGAAATAGCCAGCTGTGAGATGTT 180
    |||||
Db 142 GGCATCTCTTAAGAGAGACTTTTGAAGAAATAAAGAGAAATAGCCAGCTGTGAGATGTT 201
    |||||
QY 181 GCTAAGCAATCATCAACCTAGCTGTTTATGTATTAAGCCAGAACAGATCCTATAGGGA 240
    |||||
Db 202 GCTAAGCAATCATCAACCTAGCTGTTTATGTATTAAGCCAGAACAGATCCTATAGGGA 261
    |||||
QY 241 TTGGCACTTCTGTGTGATCTGTGTAACCCAGACTGAGTGGCTCCAGAACTTAGAAAAA 300
    |||||
Db 262 TTGGCACTTCTGTGTGATCTGTGTAACCCAGACTGAGTGGCTCCAGAACTTAGAAAAA 321
    |||||
QY 301 GCCATCCAAATTATGTATCAAAACCTGACGCAAGATGGGCTGAGAAAATTCACTGGAG 360
    |||||
Db 322 GCCATCCAAATTATGTATCAAAACCTGACGCAAGATGGGCTGAGAAAATTCACTGGAG 381
    |||||
QY 361 CCAGTGAGAAATACCCCACTGGAGAGGGAGAGAAATCAGCTGTGATGCTGGAGCCAGA 420
    |||||
Db 382 CCAGTGAGAAATACCCCACTGGAGAGGGAGAGAAATCAGCTGTGATGCTGGAGCCAGA 441
    |||||
QY 421 ATTCATAGATAGCCATCTGGGTCTTTGGCAGCAGCATTTGGGACTCTCCAGAAAGCATT 480
    |||||
Db 442 ATTCATAGATAGCCATCTGGGTCTTTGGCAGCAGCATTTGGGACTCTCTCCAGAAAGCATT 501
    |||||
QY 481 ACAGCAGAAAGTCTGTGTGTGACCTTTTCGATGAATGACAGAGAAAGGCTCAGAAACA 540
    |||||
Db 502 ACAGCAGAAAGTCTGTGTGTGACCTTTTCGATGAATGACAGAGAAAGGCTCAGAAACA 560
    |||||
QY 541 AGAGGAAAGATTGTGTGTTTATTAACCA 567
    |||||
Db 561 AGAGGAAAGATTGTGTGTTTATTAACCA 587
    |||||
```

Search completed: December 22, 2003, 23:11:34  
Job time : 1114.05 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2003, 13:59:40 ; Search time 37.4496 Seconds  
(without alignments)  
6682.695 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_57\_623

Perfect score: 567

Sequence: 1 CCTATCAGATTATCTTAAACA.....AGATTGTTTATTAAACCA 567

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents NA.\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	567	100.0	1863	US-09-482-273-28	Sequence 28, Appl
2	49.2	8.7	7218	US-08-233-463-14	Sequence 14, Appl
3	33.6	5.9	1650	US-09-220-132-103	Sequence 103, Appl
4	33.2	5.9	2443	US-08-745-934-2	Sequence 2, Appl
5	33.2	5.9	9763	US-08-973-273-1	Sequence 1, Appl
6	33	5.8	1165	US-09-023-942A-28	Sequence 28, Appl
7	33	5.8	4791	US-09-328-352-1416	Sequence 1416, Appl
8	32.6	5.7	2489	US-09-141-047-7	Sequence 7, Appl
9	32.4	5.7	3483	US-09-620-312D-658	Sequence 658, Appl
10	32.2	5.7	17949	US-09-087-465-3	Sequence 3, Appl
11	31.8	5.6	615	US-09-107-532A-1080	Sequence 1080, Appl
12	31.8	5.6	2427	US-09-134-001C-268	Sequence 268, Appl
13	31.8	5.6	11811	US-09-078-294-7	Sequence 7, Appl
14	31.6	5.6	949	US-09-221-017B-594	Sequence 594, Appl
15	31.6	5.6	28473	US-08-961-527-83	Sequence 83, Appl
16	31.2	5.5	1712	US-08-809-336A-8	Sequence 8, Appl
17	31.2	5.5	1712	US-09-688-914A-8	Sequence 8, Appl
18	31.2	5.5	1712	US-09-688-913A-8	Sequence 8, Appl
19	31.2	5.5	1712	US-09-688-916A-8	Sequence 8, Appl
20	31.2	5.5	1927	US-09-689-914A-13	Sequence 13, Appl
21	31.2	5.5	1927	US-09-689-913A-13	Sequence 13, Appl
22	31.2	5.5	1927	US-09-689-916A-13	Sequence 13, Appl
23	31.2	5.5	1954	US-08-809-336A-13	Sequence 13, Appl
24	31	5.5	1155	US-09-134-001C-1364	Sequence 1364, Appl
25	31	5.5	1497	US-09-220-132-94	Sequence 94, Appl
26	31	5.5	123025	US-09-198-452A-1	Sequence 1, Appl
27	30.8	5.4	707	US-08-998-416-1119	Sequence 1119, Appl

C	28	30.8	5.4	580073	4	US-08-545-528D-1	Sequence 1, Appl
	29	30.6	5.4	8000	4	US-09-415-784-101	Sequence 101, Appl
	30	30.6	5.4	8000	4	US-09-415-784-102	Sequence 102, Appl
	31	30.6	5.4	8000	4	US-09-415-785A-101	Sequence 101, Appl
	32	30.6	5.4	8000	4	US-09-415-785A-102	Sequence 102, Appl
	33	30.6	5.4	8000	4	US-08-944-465-101	Sequence 101, Appl
	34	30.6	5.4	8000	4	US-08-944-465-102	Sequence 102, Appl
	35	30.6	5.4	8000	4	US-09-415-868-101	Sequence 101, Appl
	36	30.6	5.4	8000	4	US-09-415-868-102	Sequence 102, Appl
	37	30.6	5.4	8000	4	US-09-415-900-101	Sequence 101, Appl
	38	30.6	5.4	8000	4	US-09-415-900-102	Sequence 102, Appl
	39	30.6	5.4	9951	4	US-09-193-707-3	Sequence 3, Appl
	40	30.6	5.4	10524	4	US-09-733-042-1	Sequence 4, Appl
	41	30.6	5.4	11282	4	US-08-484-044-10	Sequence 10, Appl
	42	30.6	5.4	11613	1	US-08-801-263A-8	Sequence 8, Appl
	43	30.6	5.4	11703	3	US-09-102-248-8	Sequence 8, Appl
	44	30.6	5.4	11703	3	US-09-367-764-8	Sequence 8, Appl
	45	30.6	5.4	11703	4		

## ALIGNMENTS

RESULT 1  
US-09-482-273-28

Sequence 28, Application US/09482273

Patent No. 6534631

GENERAL INFORMATION:

APPLICANT: Rosen et al.

FILE REFERENCE: P2030P1

CURRENT FILING DATE: 2000-01-13

EARLIER FILING DATE: 1998-07-15

EARLIER APPLICATION NUMBER: 60/092,921

EARLIER FILING DATE: 1998-07-15

EARLIER APPLICATION NUMBER: 60/092,922

EARLIER FILING DATE: 1998-07-15

EARLIER APPLICATION NUMBER: 60/092,956

NUMBER OF SEQ ID NOS: 267

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 28

LENGTH: 1863

TYPE: DNA

ORGANISM: Homo sapiens

US-09-482-273-28

Query Match 100.0%; Score 567; DB 4; Length 1863;  
Best Local Similarity 100.0%; Pred. No. 9.9e-166;

Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy	1	CCTATCAGATTATCTTAAACAAGAAACCACTGGAAAAAATGAATTCCTATCTTC	60
Db	57	CCTATCAGATTATCTTAAACAAGAAACCACTGGAAAAAATGAATTCCTATCTTC	116
Cy	61	GCATTTTCGGTGTGTTACCTTTATCCCTGTGCTTGGGAAAGCTATATGCAAGAT	120
Db	117	GCATTTTCGGTGTGTTACCTTTATCCCTGTGCTTGGGAAAGCTATATGCAAGAT	176
Cy	121	GGCATCTCTAAGAGACCTTTTGAAGAATAAAGAAATAGCAGCTGGAGATGTT	180
Db	177	GGCATCTCTAAGAGACCTTTTGAAGAATAAAGAAATAGCAGCTGGAGATGTT	236
Cy	181	GCTAAAGCAATCATCAACTAGTGTATATGTAAGCCAGAACAGATCCTATAGCGA	240
Db	237	GCTAAAGCAATCATCAACTAGTGTATATGTAAGCCAGAACAGATCCTATAGCGA	296
Cy	241	TTGGCACTTCTGTTGATCTGTTGACCCAGATGAGTGGCTCAAGAACTTGAAGAA	300
Db	297	TTGGCACTTCTGTTGATCTGTTGACCCAGATGAGTGGCTCAAGAACTTGAAGAA	356

OY 301 GCATCCAAATATATGACCAAAACCTGACGACGAAGATGGCTGAGAAAGTTCACTGAG 360  
DB 357 GCATCCAAATATATGACCAAAACCTGACGACGAAGATGGCTGAGAAAGTTCACTGAG 416  
OY 361 CGAGTGAATATCCCTGAGGAGGAGGAGGAAGATCACTGTGATCTGAGAGCAAGA 420  
DB 417 CCAGTGAATATCCCTGAGGAGGAGGAGGAAGATCACTGTGATCTGAGAGCAAGA 476  
OY 421 ATTCAATAGATGACCATCTGGGCTCTTGACGACGACATTTGGGACCTCCGAAAGCAAT 480  
DB 477 ATTCAATAGATGACCATCTGGGCTCTTGACGACGACATTTGGGACCTCCGAAAGCAAT 536  
OY 481 ACAGAGAAGTCTGCTGATGACCTCTTTCATGATGACGAGAAAGGCTCAGAAACA 540  
DB 537 ACAGAGAAGTCTGCTGATGACCTCTTTCATGATGACGAGAAAGGCTCAGAAACA 596  
OY 541 AGAGGAAGATTTGTTTATTAACCA 567  
DB 597 AGAGGAAGATTTGTTTATTAACCA 623

## RESULT 2

US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114, IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-Fls  
; US-08-232-463-14

Query Match 8.7%; Score 49.2; DB 1; Length 7218;  
Best Local Similarity 8.1%; Pred. No. 4,7e-05;  
Matches 36; Conservative 215; Mismatches 193; Indels 0; Gaps 0;

OY 91 CTGCTCTGGGAAAGCTATATGCAAGATGCATCTTAAGAGACTTTGAAGATA 150  
DB 1471 CTATGCAAGTATGTTAAAGATAGAGATTTGGTACRRRRRRRRRRRRRRRR 1412  
OY 151 AAGAGAAATAGCCAGCTGTGAGATGTTCTTAAGCAATCATCACTAGCTTAT 210  
DB 1411 RRR 1352  
OY 211 GGTAAAGCCAGAAACATCTATGACGATTTGCCTTCTGTGATCTGTTGACC 270  
DB 1351 RRR 1292  
OY 271 AGACTGAGTGGCTCAAGAACTAGAAAAGCCATCAAAATATGATCAAAAGCTGAC 330  
DB 1291 RRR 1232  
OY 331 CAAGATGGCTGAGAAAGTTACCTGAGACCACTGAGAAATACCCACTGGAGAGGGA 390  
DB 1231 RRR 1172  
OY 391 GAAGATCAGCTGTGATGCTGAGCCAGAAATTAATAGATGACCTCTGCTTGGC 450  
DB 1171 RRR 1112  
OY 451 AGCAGATTTGGACTCTCCAGAAAGCATTAAGCAGAGATTTGTTGTTGACTCTTTC 510  
DB 1111 RRR 1052  
OY 511 GATGACTGAGAGAGGCTCA 534  
DB 1051 GACTGAGCAAGCTCGAATTA 1028

## RESULT 3

US-09-220-132-103  
; Sequence 103, Application US/09220132  
; Patent No. 6506607  
; GENERAL INFORMATION:  
; APPLICANT: Shyjan, Andrew W.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT  
; FILE REFERENCE: 07334-074001  
; CURRENT APPLICATION NUMBER: US/09/220,132  
; CURRENT FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: US 60/079,303  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: US 60/068,821  
; PRIOR FILING DATE: 1997-12-24  
; NUMBER OF SEQ ID NOS: 191  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 103  
; LENGTH: 1650  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-220-132-103

Query Match 5.9%; Score 33.6; DB 4; Length 1650;  
Best Local Similarity 61.4%; Pred. No. 1.5;  
Matches 54; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

OY 41 AATGAATTCCTATCTTGGCAATTTTGGTGTGTCACCTTATCCCTGTGCTG 100  
DB 1313 AACTGAATGATTTGCTCCACATTAATCGGTGGCTGATATTAATGCTCATGATG 1372  
OY 101 GGAAGCTATATGCAAGATGCAATGCTC 128  
DB 1373 AGAAGATATGCAAAATTAATATCTC 1400

## RESULT 4

US-08-745-934-2/c  
; Sequence 2, Application US/08745934

Patent No. 5861496  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Hawkins, Phillip R.  
TITLE OF INVENTION: HUMAN SQUALENE EPOXIDASE  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/745,934  
FILING DATE: Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0151 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 855-0555  
TELEFAX: (415) 845-4166  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2443 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE: Consensus  
US-08-745-934-2

Query Match 5.9%; Score 33.2; DB 2; Length 2443;  
Best Local Similarity 50.0%; Pred. No. 2.5;  
Matches 83; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 170 GTGAGATGTTGCTAAAGCAATCATCAACCTAGCTGTTATGTTAAAGCCGAGAACAGAT 229  
Db 741 GTATGTAAGTCCCGAATCTTATTAATAAGTGAAGTGAATGCCAGAAAGTCC 682

QY 230 CCTATGAGCAATGGCACTTCTGTTGATCTGTTGAGACCACTGAGTGGCTCCAGA 289  
Db 681 CACATGTTCCAAAGCTTCTTGTGTCGACCGGTCAAGCGGAGATTATCTGCTGAGTG 622

QY 290 ACCGAGAAAAGCAATCAATTATGACCAAAACGCGAGCAGAGA 335  
Db 621 GGTTAAGTTCTCCCAATGAGGTACCAAGAGACGAGAGA 576

RESULT 5  
US-08-973-273-1/c  
Sequence 1, Application US/08973273  
Patent No. 6140085  
GENERAL INFORMATION:  
APPLICANT: Dean, Caroline  
APPLICANT: MacKnight, Richard C  
APPLICANT: Bancroft, Ian  
APPLICANT: Lister, Clare K  
TITLE OF INVENTION: Genetic Control of Flowering  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vandermere P.C.

STREET: 1100 No. 6140085th Giebe Road, 8th Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: United States of America  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/973,273  
FILING DATE: 01-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB96/01332  
FILING DATE: 03-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9511196.9  
FILING DATE: 02-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ms Mary J Wilson  
REGISTRATION NUMBER: 32,955  
REFERENCE/DOCKET NUMBER: 620-29  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9763 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHEetical: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Arabidopsis thaliana  
STRAIN: Columbia  
US-08-973-273-1

Query Match 5.9%; Score 33.2; DB 3; Length 9763;  
Best Local Similarity 61.6%; Pred. No. 5;  
Matches 53; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 131 AGAGACTTTTGAAGAAATTAAGAAATATGACCTGTGAGATGTTGCTAAAGCA 190  
Db 8640 AGAGAGTTCTGATCATCATGATTAACCACTACTGTTGAGATGTTGATATATCAT 8581

QY 191 TCATCAACTAGCTGTTATGTTAA 216  
Db 8580 TCATGAGCACTCCACGCGGTATA 8555

RESULT 6  
US-09-023-942A-28/c  
Sequence 28, Application US/09023942A  
Patent No. 6479274  
GENERAL INFORMATION:  
APPLICANT: (US only) ANTALIS Toni Marie and HOOBER John David  
TITLE OF INVENTION: NOVEL MOLECULES  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 GARDEN CITY PLAZA  
CITY: GARDEN CITY  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/023,942A  
;; FILING DATE: 13-FEB-1998  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: POS101/97  
;; FILING DATE: 13-FEB-1997  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PP0422/97  
;; FILING DATE: 18-NOV-1997  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: International PCT Application  
;; FILING DATE: 13-FEB-1998  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: DIGICLO, FRANK S  
;; REGISTRATION NUMBER: 31,346  
;; REFERENCE/DOCKET NUMBER: 11168  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (516) 742 4363  
;; TELEFAX: (516) 742 4366  
;; TELEX: 230 901 SANS UR  
;; INFORMATION FOR SEQ ID NO: 28:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1165 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA  
;; US-09-023-942A-28  
  
Query Match  
Best Local Similarity 5.8%; Score 33; DB 4; Length 1165;  
Matches 72; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
  
Qy 280 GCTCCAGAACTTGAAGAAAGCCATCCAAATTATGTACCAAAACCTGCACCAAGATGCG 339  
Db 1126 GATTACAATATCTTAAATAATCATTAATTATATACAAAAGCCGCGGCGCCGCG 1067  
Qy 340 CTGAGAAAGTTCACTGAGCGGAGTGAATACCCCACTGGGAGAGGGAGGAAGATCA 399  
Db 1066 GCGGGGCGGAGTGTGGGGGCGGAGATGAGGCGTGGCGGGGCGCTGATGCCCTTG 1007  
Qy 400 GCTGTGATGCTGAGGCC 416  
Db 1006 GAGGCGGGGCTGAGGC 990  
  
RESULT 7  
US-09-328-352-1416  
; Sequence 1416, Application US/09328352  
; Patent No. 6562358  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 1416  
; LENGTH: 4791  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-1416  
  
Query Match  
Best Local Similarity 5.8%; Score 33; DB 4; Length 4791;  
Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;  
  
Qy 124 ATCTCTAAGAGACTTTGAAAGAAATTAAGAAGAAATAGCCAGCTGTGAGAGTGTCT 183  
Db 1528 ATCATTTGGGTTTGTAGATGACAAAGCTGAGCCATTAGACGTTATGATAGATACC 1587

Qy 184 AAGCAATCATCAACCTAGCTGTTATGTATGTAAGCCAGAGACAGATCTATGA 236  
Db 1588 CAAGGAACTTAATTAAGCTATTGATCAGAAATGCCATACAGAACTTATGA 1640  
  
RESULT 8  
US-09-141-047-7  
; Sequence 7, Application US/09141047A  
; Patent No. 6043085  
; GENERAL INFORMATION:  
; APPLICANT: Walker, David H.  
; TITLE OF INVENTION: Ehrlichia canis 120-kDa Immunodominant Antigenic  
; FILE REFERENCE: D6143  
; CURRENT APPLICATION NUMBER: US/09/141,047A  
; CURRENT FILING DATE: 1998-08-27  
; NUMBER OF SEQ ID NOS: 15  
; SEQ ID NO 7  
; LENGTH: 2489  
; TYPE: DNA  
; ORGANISM: Ehrlichia canis  
; FEATURE:  
; LOCATION: -340..2149  
; OTHER INFORMATION: Nucleotide sequence of gene encoding 120kDa  
; OTHER INFORMATION: Immunoreactive protein  
US-09-141-047-7  
  
Query Match  
Best Local Similarity 5.7%; Score 32.6; DB 3; Length 2489;  
Matches 86; Conservative 0; Mismatches 89; Indels 0; Gaps 0;  
  
Qy 15 TTACAGAAAGCAACTGGAAGAAAGAAATGAATTCCTATCTTCGCAATTTTCGCTG 74  
Db 538 TAAAGAAAGAAAGCTCTCTGAAGTTAAAGCGAAGATTTCACCTGCTGATGAGTAG 597  
Qy 75 TGTCACCTTTATCCCTGTCTCTGGAAAGCTATATGCAAGATGGCATCTTAAGAG 134  
Db 598 TGTGAACATTATCAAGTGAAGTGGGAAAAAGTATCGAAACTATGTAAGAGGAAAG 657  
Qy 135 GACTTTGAAGAAATTAAGAAGAAATAGCCAGCTGTGAGATGTTGCTTAAGCA 189  
Db 658 TACTCTGAAGTTAAAGCAGAAAGATTGCAACTGCTGTAGATGGTATAGAA 712  
  
RESULT 9  
US-09-620-312D-658  
; Sequence 658, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Weinman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yundong  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Dmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: PC\_Fl\_genes Version 1.0  
SEQ ID NO: 658  
LENGTH: 3483  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (94)..(2064)  
US-09-620-312D-658

Query Match 5.7%; Score 32.4; DB 4; Length 3483;  
Best Local Similarity 58.2%; Pred. No. 5.2;  
Matches 57; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 101 GGAAGCTATATGCAAGATGCGATCTTAAGAGGACTTTGAAGAATAAAGAGAAA 160  
DB 335 GTAAAGTTAGCCATCATGAGATTTGTAGCATTTTCAAGAGAGTGAAGAGAAC 394  
QY 161 TAGCCAGCTGTGAGATGTTGCTAAAGCATCATCAAC 198  
DB 395 TTGAAGCATTAAGGAGAGATGTTCAAGCAATGAGCAAC 432

RESULT 10  
US-09-087-465-3/c  
Sequence 3, Application US/09087465A  
Patent No. 6160092  
GENERAL INFORMATION:  
APPLICANT: Vinkemeier, Uwe  
APPLICANT: Chen, Xiaomin  
APPLICANT: Darrell Jr., James E  
APPLICANT: Kuriyan, John  
TITLE OF INVENTION: A CRYSTAL OF THE CORE PORTION OF A STAT AND METHODS OF  
FILE REFERENCE: 600-1-229  
CURRENT APPLICATION NUMBER: US/09/087,465A  
CURRENT FILING DATE: 1998-05-29  
NUMBER OF SEQ ID NOS: 37  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 3  
LENGTH: 17949  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-087-465-3

Query Match 5.7%; Score 32.2; DB 3; Length 17949;  
Best Local Similarity 54.7%; Pred. No. 14;  
Matches 64; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
QY 5 TCAGATTATCTTAAACAAGAAACCACTGAGAAAAAATGAAATTCCTTATCTTGGCAT 64  
DB 15045 TGAAGACTGTCTCAAAAAAATAAATAAATAAATAAATAAATTCATCCCGACGCA 14986  
QY 65 TTTTGGTGTGTTCACTTTTATCCCTGTGCTGTGAGAAAGTATATATGCAAGATG 121  
DB 14985 AGTGCCATGGCTCATCTTCAATCCCGCAATTTGGAGAGCAAGAGAGACTG 14929

RESULT 11  
US-09-107-532A-1080  
Sequence 1080, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ardiello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 1080:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 615 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...615  
SEQUENCE DESCRIPTION: SEQ ID NO: 1080:  
US-09-107-532A-1080  
Query Match 5.6%; Score 31.8; DB 4; Length 615;  
Best Local Similarity 49.7%; Pred. No. 3.3;  
Matches 81; Conservative 0; Mismatches 82; Indels 0; Gaps 0;  
QY 128 CTAAGAGACTTTTGAAGAAATAAAGAGAAATAGCCAGCTGTGAGATGTTGCTAAAG 187  
DB 332 CTCAGGGAATCTTGAAAGAAAAACAAGATGATGAGAGAAATTAATTCGGCTGAAA 391  
QY 188 CAATCATCAACTAGCTGTTTATGTTAAGCCAGAAACAGATCTTATGAGCGATTGGCAG 247  
DB 392 AAGCTAAAGAGAGCTCTGATTAACGATCCACTAAGAAAGCTTGGAAAAATTAGTGC 451  
QY 248 TTTGTTGATATGTTGAGCCAGCTGAGCTGCTCCAGAA 290  
DB 452 TGAACATTGAAATGTTCTTTCGCAAGTTAGTGGCTACGAA 494  
RESULT 12  
US-09-134-001C-268  
Sequence 268, Application US/09134001C  
Patent No. 6580370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779

;; PRIOR FILING DATE: 1997-08-14  
;; NUMBER OF SEQ ID NOS: 5674  
;; SEQ ID NO 268  
;; LENGTH: 2427  
;; TYPE: DNA  
;; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-268

Query Match 5.6%; Score 31.8; DB 4; Length 2427;  
Best Local Similarity 52.7%; Pred. No. 6.6; Indels 0; Gaps 0;  
Matches 69; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 30 ACTGGAAAAAATGAATTCCTTATCTTCGCATTTTCGGTGTGTCACCTTTATC 89  
DB 2265 ACATGTAATAATATGTAATATGTCATCTCGTAATTCGCTGTATGATCTTAT 2324  
QY 90 CCTGTCTCTGGAAAGCTATATGCAAGATGGCATCTCTAAGAGACTTTGAAGAA 149  
DB 2325 CGAGTCATCTGTAAGATGCGAGTTAGCGCTATATCAATGATTAATTAAGATT 2384  
QY 150 AAAAGAGAAA 160  
DB 2385 AAAACAGAAA 2395

RESULT 13  
US-09-078-294-7/c  
; Sequence 7, Application US/09078294  
; Patent No. 6265211  
; GENERAL INFORMATION:  
; APPLICANT: Choo, Kong-Hong Andy  
; APPLICANT: Du Sart, Desiree  
; APPLICANT: Cancilla, Michael R.  
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE  
; FILE REFERENCE: Davies Col  
; CURRENT APPLICATION NUMBER: US/09/078,294  
; CURRENT FILING DATE: 1998-05-13  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 11811  
; TYPE: DNA  
; ORGANISM: BAC-F2 config 3  
US-09-078-294-7

Query Match 5.6%; Score 31.8; DB 3; Length 11811;  
Best Local Similarity 56.1%; Pred. No. 15; Indels 0; Gaps 0;  
Matches 60; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 16 TAACAGAAAAACAACCTGGAAAAAATGAATTCCTTATCTTCGCATTTTCGGTGT 75  
DB 6086 TAAAAAATAAATAATAGTAAAAACAAACAAACAAACCTGCTGTGATGTGGC 6027  
QY 76 GTTCACTTTTATCCCTGTCTCTGGAAAGCTATATGCAAGATG 122  
DB 6026 TCACCCCTGTATCCCAACATTTGGAGAGCCAGGTGATGATCG 5980

RESULT 14  
US-09-221-017B-594  
; Sequence 594, Application US/09221017B  
; Patent No. 6444799  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Bruce C.  
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
; NUMBER OF SEQUENCES: 1120  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/221,017B  
FILING DATE: 23-DEC-1998

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP1182  
FILING DATE: 31-DEC-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP1546  
FILING DATE: 30-JAN-1998

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP2911  
FILING DATE: 09-APR-1998

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU98/01023  
FILING DATE: 10-DEC-1998

ATTORNEY/AGENT INFORMATION:  
NAME: Montoy, Gladys H  
REGISTRATION NUMBER: 32,430

REFERENCE/DOCKET NUMBER: 27340-20021.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-813-5600  
TELEFAX: 650-494-0792

TELEX: 706141  
INFORMATION FOR SEQ ID NO: 594:

SEQUENCE CHARACTERISTICS:  
LENGTH: 949 base pairs  
TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: UNKNOWN

ORIGINAL SOURCE:  
ORGANISM: PORYPHYROMONAS GINGIVALIS  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1...949

US-09-221-017B-594

Query Match 5.6%; Score 31.6; DB 4; Length 949;  
Best Local Similarity 60.5%; Pred. No. 4.8; Indels 34; Gaps 0;  
Matches 52; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 349 GTTCACTGGAGCCAGTGAATATCCCACTGGAGAGGGGAGAGATCAGTGTGATG 408  
DB 251 GTACACGCGGTGTGAATGCTATGACATCCCTGCTGACGCGTCAAGATACGCCGTTCTG 310  
QY 409 CTGAGCCAGAGATTCATTAAGATAGC 434  
DB 311 CTTGGCCATCATATAGAGAGAGC 336

RESULT 15  
US-08-961-527-83/c  
; Sequence 83, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850



```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961.527

FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ. ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 28473 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

```

Query Match	5.6%	Score	31.6	DB	4	Length	28473
Best Local Similarity	54.2%	Pred. No.	27				
Matches	64	Conservative	0	Mismatches	54	Indels	0
						Gaps	0

QY	163	24092	24093	24094	24095	24096	24097	24098	24099	24100	24101	24102	24103	24104	24105	24106	24107	24108	24109	24110	24111	24112	24113	24114	24115	24116	24117	24118	24119	24120	24121	24122	24123	24124	24125	24126	24127	24128	24129	24130	24131	24132	24133	24134	24135	24136	24137	24138	24139	24140	24141	24142	24143	24144	24145	24146	24147	24148	24149	24150	24151	24152	24153	24154	24155	24156	24157	24158	24159	24160	24161	24162	24163	24164	24165	24166	24167	24168	24169	24170	24171	24172	24173	24174	24175	24176	24177	24178	24179	24180	24181	24182	24183	24184	24185	24186	24187	24188	24189	24190	24191	24192	24193	24194	24195	24196	24197	24198	24199	24200	24201	24202	24203	24204	24205	24206	24207	24208	24209	24210	24211	24212	24213	24214	24215	24216	24217	24218	24219	24220	24221	24222	24223	24224	24225	24226	24227	24228	24229	24230	24231	24232	24233	24234	24235	24236	24237	24238	24239	24240	24241	24242	24243	24244	24245	24246	24247	24248	24249	24250	24251	24252	24253	24254	24255	24256	24257	24258	24259	24260	24261	24262	24263	24264	24265	24266	24267	24268	24269	24270	24271	24272	24273	24274	24275	24276	24277	24278	24279	24280	24281	24282	24283	24284	24285	24286	24287	24288	24289	24290	24291	24292	24293	24294	24295	24296	24297	24298	24299	24300	24301	24302	24303	24304	24305	24306	24307	24308	24309	24310	24311	24312	24313	24314	24315	24316	24317	24318	24319	24320	24321	24322	24323	24324	24325	24326	24327	24328	24329	24330	24331	24332	24333	24334	24335	24336	24337	24338	24339	24340	24341	24342	24343	24344	24345	24346	24347	24348	24349	24350	24351	24352	24353	24354	24355	24356	24357	24358	24359	24360	24361	24362	24363	24364	24365	24366	24367	24368	24369	24370	24371	24372	24373	24374	24375	24376	24377	24378	24379	24380	24381	24382	24383	24384	24385	24386	24387	24388	24389	24390	24391	24392	24393	24394	24395	24396	24397	24398	24399	24400	24401	24402	24403	24404	24405	24406	24407	24408	24409	24410	24411	24412	24413	24414	24415	24416	24417	24418	24419	24420	24421	24422	24423	24424	24425	24426	24427	24428	24429	24430	24431	24432	24433	24434	24435	24436	24437	24438	24439	24440	24441	24442	24443	24444	24445	24446	24447	24448	24449	24450	24451	24452	24453	24454	24455	24456	24457	24458	24459	24460	24461</
----	-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	---------

Search completed: December 22, 2003, 23:17:23  
Job time : 39.4496 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 22, 2003, 14:49:05 ; Search time 177.393 Seconds  
(without alignments)  
10650.710 Million cell updates/sec

Title: US-09-745-763-35\_COPY\_57\_623

Perfect score: 567  
Sequence: 1 CCGATCAGATTATCTTACCA.....AGATTGTTGTTATACCA 567

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2211978 seqs, 1666101734 residues

Total number of hits satisfying chosen parameters: 4423956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	567	100.0	1851	9	US-09-745-763-35
2	567	100.0	1863	11	US-09-984-271-28
3	399.6	70.5	1778	10	US-09-917-800A-505
4	247.4	43.6	427	10	US-09-833-381-1930
5	60	10.6	60	13	US-09-908-975-8960
6	41	7.2	63	13	US-09-908-975-1202
7	39.4	6.9	1363	11	US-09-918-995-30889
8	39.4	6.9	1363	15	US-10-232-484-5
9	39.4	6.9	2319	14	US-10-044-090-579
10	39.4	6.9	3322	13	US-09-814-353-20277
11	39	6.9	665	13	US-10-027-632-101927
12	39	6.9	665	13	US-10-027-632-101927
13	36.6	6.5	786431	14	US-10-027-632-101927
14	36.2	6.4	505	15	US-10-066-543-2072
15	36.2	6.4	1767	11	US-09-764-891-9681

C 16	36.2	6.4	715517	13	US-10-027-632-53712	Sequence 53712, A
C 17	36.2	6.4	715517	14	US-10-027-632-53712	Sequence 53712, A
C 18	36	6.3	17967	13	US-10-311-455-987	Sequence 987, App
C 19	35.2	6.2	513509	11	US-09-754-853A-4	Sequence 4, Appl1
C 20	35	6.2	390	11	US-09-918-995-35565	Sequence 35565, A
C 21	35	6.2	1136	13	US-10-027-632-249053	Sequence 249053, A
C 22	35	6.2	1136	14	US-10-027-632-249053	Sequence 249053, A
C 23	35	6.2	2048	13	US-10-094-749-1415	Sequence 1415, Ap
C 24	34.6	6.1	888	13	US-10-027-632-122347	Sequence 122347, A
C 25	34.6	6.1	888	14	US-10-027-632-122347	Sequence 122347, A
C 26	34.6	6.1	2155	13	US-10-027-632-252101	Sequence 252101, A
C 27	34.6	6.1	2155	14	US-10-027-632-252101	Sequence 252101, A
C 28	34.2	6.0	514	13	US-09-814-353-14775	Sequence 14775, A
C 29	34.2	6.0	552	13	US-09-814-353-2045	Sequence 2045, Ap
C 30	34.2	6.0	552	13	US-09-814-353-8391	Sequence 8391, Ap
C 31	34.2	6.0	708	13	US-10-027-632-28380	Sequence 28380, A
C 32	34.2	6.0	708	14	US-10-027-632-28380	Sequence 28380, A
C 33	34.2	6.0	5781	10	US-09-764-868-1374	Sequence 1374, Ap
C 34	34	6.0	783	13	US-10-027-632-62140	Sequence 62140, A
C 35	34	6.0	783	14	US-10-027-632-62140	Sequence 62140, A
C 36	34	6.0	882	9	US-09-815-242-4206	Sequence 4206, Ap
C 37	34	6.0	900	9	US-09-815-242-8477	Sequence 8477, Ap
C 38	34	6.0	5049	11	US-09-764-891-6588	Sequence 6588, Ap
C 39	34	6.0	5049	11	US-09-764-891-6589	Sequence 6589, Ap
C 40	34	6.0	5049	15	US-10-091-572-460	Sequence 460, App
C 41	34	6.0	5049	15	US-10-091-572-461	Sequence 461, App
C 42	34	6.0	34450	13	US-10-293-832-3	Sequence 3, Appl1
C 43	34	6.0	34757	13	US-10-293-832-22	Sequence 22, Appl1
C 44	33.6	5.9	789	13	US-10-027-632-123392	Sequence 123392, A
C 45	33.6	5.9	789	14	US-10-027-632-123392	Sequence 123392, A

#### ALIGNMENTS

##### RESULT 1

US-09-745-763-35  
Sequence 35, Application US/09745763  
Patent No. US20020065394A1

##### GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth  
McCoy, John M.  
LaValle, Edward R.  
Collins-Racie, Lisa A.  
Evans, Cheryl  
Merberg, David  
Treacy, Maurice  
Spaulding, Vikki  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM  
NUMBER OF SEQUENCES: 219  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140

##### COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/745,763  
FILING DATE: 18-Jun-2000  
CLASSIFICATION: <Unknown>

##### ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.  
REGISTRATION NUMBER: 41,323  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8284  
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1851 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 35:  
US-09-745-763-35

Query Match 100.0%; Score 567; DB 9; Length 1851;  
Best Local Similarity 100.0%; Pred. No. 9,2e-165;  
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CCTATCGATTATCTTAAACAAGAAACCACTGGAAGAAATGAAATCTTATCTTC 60
DB 57 CCTATCGATTATCTTAAACAAGAAACCACTGGAAGAAATGAAATCTTATCTTC 116
QY 61 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
DB 117 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 176
QY 121 GGCACTCTTAAGAGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGT 180
DB 177 GGCACTCTTAAGAGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGT 236
QY 181 GCTTAAGCAATCATCAACCTGCTGTTATGTTAAAGCCAGAAACAGATCTTATGAGCGA 240
DB 237 GCTTAAGCAATCATCAACCTGCTGTTATGTTAAAGCCAGAAACAGATCTTATGAGCGA 296
QY 241 TTGGCACTTCTGTTGATCTGTGAGACCCAGACTGAGTGCTCCAGAAACCTAGAAAAA 300
DB 297 TTGGCACTTCTGTTGATCTGTGAGACCCAGACTGAGTGCTCCAGAAACCTAGAAAAA 356
QY 301 GCCATCCAAATTATGTACCAAAACCTGACGAGAAATGGCTGAGAAAGTTCACTTGAG 360
DB 357 GCCATCCAAATTATGTACCAAAACCTGACGAGAAATGGCTGAGAAAGTTCACTTGAG 416
QY 361 CCAATGGAATATCCCACTGGGAGAGGGGAGAAAGATAGCTGTGATCTGAGCCCAAGA 420
DB 417 CCAATGGAATATCCCACTGGGAGAGGGGAGAAAGATAGCTGTGATCTGAGCCCAAGA 476
QY 421 ATTCAATAGATAGCATCTGGGTCTTTGGCAGACAGATTGGGACTCTCCAGAAAGCAT 480
DB 477 ATTCAATAGATAGCATCTGGGTCTTTGGCAGACAGATTGGGACTCTCCAGAAAGCAT 536
QY 481 ACAGCAGAAAGTTCTGTGTGATGACCTCTTTCATGATGAATGCAAGAAAGGCTCAGAA 540
DB 537 ACAGCAGAAAGTTCTGTGTGATGACCTCTTTCATGATGAATGCAAGAAAGGCTCAGAA 596
QY 541 AGAGGGAAGATTTGTTTATTAACCA 567
DB 597 AGAGGGAAGATTTGTTTATTAACCA 623
```

RESULT 2  
US-09-984-271-28  
Sequence 28, Application US/09984271  
Publication No. US20030040088A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 71 Human Secreted Proteins  
FILE REFERENCE: P2030P1  
CURRENT APPLICATION NUMBER: US/09/984,271  
CURRENT FILING DATE: 2001-10-29  
PRIOR APPLICATION NUMBER: 09/482,273  
PRIOR FILING DATE: 2000-01-13  
PRIOR APPLICATION NUMBER: PCT/US99/15849  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: 60/092,921  
PRIOR FILING DATE: 1998-07-15  
PRIOR APPLICATION NUMBER: 60/092,922  
PRIOR FILING DATE: 1998-07-15

PRIOR APPLICATION NUMBER: 60/092,956  
PRIOR FILING DATE: 1998-07-15  
NUMBER OF SEQ ID NOS: 267  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 28  
LENGTH: 1863  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-984-271-28

Query Match 100.0%; Score 567; DB 11; Length 1863;  
Best Local Similarity 100.0%; Pred. No. 9,2e-165;  
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CCTATCGATTATCTTAAACAAGAAACCACTGGAAGAAATGAAATCTTATCTTC 60
DB 57 CCTATCGATTATCTTAAACAAGAAACCACTGGAAGAAATGAAATCTTATCTTC 116
QY 61 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 120
DB 117 GCATTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTCTGGGAAAGCTATATGCAAGAT 176
QY 121 GGCACTCTTAAGAGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGT 180
DB 177 GGCACTCTTAAGAGACTTTTGAAGAAATAAAGAAATAGCCAGCTGTGAGATGT 236
QY 181 GCTTAAGCAATCATCAACCTGCTGTTATGTTAAAGCCAGAAACAGATCTTATGAGCGA 240
DB 237 GCTTAAGCAATCATCAACCTGCTGTTATGTTAAAGCCAGAAACAGATCTTATGAGCGA 296
QY 241 TTGGCACTTCTGTTGATCTGTGAGACCCAGACTGAGTGCTCCAGAAACCTAGAAAAA 300
DB 297 TTGGCACTTCTGTTGATCTGTGAGACCCAGACTGAGTGCTCCAGAAACCTAGAAAAA 356
QY 301 GCCATCCAAATTATGTACCAAAACCTGACGAGAAATGGCTGAGAAAGTTCACTTGAG 360
DB 357 GCCATCCAAATTATGTACCAAAACCTGACGAGAAATGGCTGAGAAAGTTCACTTGAG 416
QY 361 CCAATGGAATATCCCACTGGGAGAGGGGAGAAAGATAGCTGTGATCTGAGCCCAAGA 420
DB 417 CCAATGGAATATCCCACTGGGAGAGGGGAGAAAGATAGCTGTGATCTGAGCCCAAGA 476
QY 421 ATTCAATAGATAGCATCTGGGTCTTTGGCAGACAGATTGGGACTCTCCAGAAAGCAT 480
DB 477 ATTCAATAGATAGCATCTGGGTCTTTGGCAGACAGATTGGGACTCTCCAGAAAGCAT 536
QY 481 ACAGCAGAAAGTTCTGTGTGATGACCTCTTTCATGATGAATGCAAGAAAGGCTCAGAA 540
DB 537 ACAGCAGAAAGTTCTGTGTGATGACCTCTTTCATGATGAATGCAAGAAAGGCTCAGAA 596
QY 541 AGAGGGAAGATTTGTTTATTAACCA 567
DB 597 AGAGGGAAGATTTGTTTATTAACCA 623
```

RESULT 3  
US-09-917-800A-505  
Sequence 505, Application US/09917800A  
Patent No. US20020119462A1  
GENERAL INFORMATION:  
APPLICANT: Mendrick, Donna  
APPLICANT: Porter, Mark  
APPLICANT: Johnson, Kory  
APPLICANT: Casale, Arthur  
APPLICANT: Elashoff, Michael  
APPLICANT: Gene Logic, Inc.  
TITLE OF INVENTION: Molecular Toxicology Modeling  
FILE REFERENCE: 44921-5038-US  
CURRENT APPLICATION NUMBER: US/09/917,800A  
CURRENT FILING DATE: 2001-07-31  
PRIOR APPLICATION NUMBER: US 60/222,040  
PRIOR FILING DATE: 2000-07-31  
PRIOR APPLICATION NUMBER: US 60/222,880

```

; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/230,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/230,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/232,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/235,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/237,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/238,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 505
; LENGTH: 1778
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AF097723
US-09-917-800A-505

Query Match
Best Local Similarity 70.5%; Score 399.6; DB 10; Length 1778;
Matches 456; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 17 AACAGAAAAACAATGAAAAAATGAAATTCCTTATCTTGCAATTTTGGGTG 76
DB 95 AGCAAGAAAAAGAACTAGAGGACATGAGTTCCTTTCTTCTGTTGCTGTG 154
QY 77 TTCACTTTTATCCCTGTGCTCTGGAAGCTATATGCAAGATGCGATCTTAAGGA 136
DB 155 TTCACTTTTCTCTGCGCTCTGGAAGCTATATGCAAGATGCGATCTTAAGGA 214
QY 137 CTTTGAAGAAATTAAGAAATAGCAAGCTGTGAGATGTTGCTAAGACATCATCA 196
DB 215 CATTTCAAGAAATTAAGAAATAGCAAGCTGTGAGATGTTGCTAAGACATCATCA 274
QY 197 ACCTAGCTTTTATGTTAAAGCCGAGAACAGATCTTAAGCAATTTGCTGTTG 256
DB 275 ACCTGCTGTTTATGAAATATACAGAACCGGTGTATGAGCGTTTGGACTTACTG 334
QY 257 ATACTGTTGACCCAGCTAGTGGCTTCAAGAACTTAAGAAAAAGCATCAATTATGT 316
DB 335 ATACTGTTGACCCAGCTAGTGGCTTCAAGAACTTAAGAAAAAGCATCAATTATGT 394
QY 317 ACCAAAACTGACAGAGATGGGCTGAGAAAGTTCACTGAGCCAGTGAATACCC 376
DB 395 ACCAAAACTGACAGAGATGGGCTGAGAAAGTTCACTGAGCCAGTGAATACCC 454
QY 377 ACTGGAGAGAGAGAGAAATCACTGTGATGCTGAGCCAAATTCATTAAGATACCA 436
DB 455 ACTGGAGAGAGAGAGAAATCACTGTGATGCTGAGCCAAATTCATTAAGATACCA 514
QY 437 TCCTGGGCTTTGGCAGCAGATTGGGACTCTTCAAGAAAGCATTAACGCAAGTTTGG 496
DB 515 TTTTAGGCTTTGGCAGCAGATTGGGACTCTTCAAGAAAGCATTAACGCAAGTTTGG 574
QY 497 TGGGAGCTCTTTGAGATGAGCTGAGAGAGGCTCAGAGCAAGAGGAGAGATTTGG 556
DB 575 TGGGAGCTCTTTGAGATGAGCTGAGAGAGGCTCAGAGCAAGAGGAGAGATTTGG 634
QY 557 TTTATTAACA 566
DB 635 TTTATTAACA 644

RESULT 4
US-09-833-381-1930
; Sequence 1930, Application US/09833381
```

```

; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: PatSeq for Windows Version 3.0
; SEQ ID NO 1930
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(427)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1930

Query Match
Best Local Similarity 43.6%; Score 247.4; DB 10; Length 427;
Matches 269; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 75 TGTTCACCTTTATCCCTGTGCTCTGGAAGCTATATGCAAGATGCGATCTTAAGAG 134
DB 117 TGTTCACCTTTATCCCTGTGCTCTGGAAGCTATATGCAAGATGCGATCTTAAGAG 176
QY 135 GACTTTGAGAAATTAAGAAATAGCAGCTGTGAGATGTTGCTAAGCAATCAT 194
DB 177 ACATTTGAGAAATTAAGAAATAGCAGCTGTGAGATGTTGCTAAGCAATCAT 236
QY 195 CAACCTAGCTTTATGTTAAAGCCAGAACAGATCTTATGAGCATTTGCTTCTGCT 254
DB 237 CAACCTAGCTTTATGTTAAAGCCAGAACAGATCTTATGAGCATTTGCTTCTGCT 296
QY 255 TGATAGCTTTGAGCCAGAGCTGAGGCTCCAGAAAGCTAAGAAAAAGCATCAATAT 314
DB 297 TGATAGCTTTGAGCCAGAGCTGAGGCTCCAGAAAGCTAAGAAAAAGCATCAATAT 356
QY 315 GTACCAAAACCTGACAGCAAGATGGCTGAGAAAGTTCACTGAGGAGCAATATAC 374
DB 357 GTACCAAAACCTGACAGCAAGATGGCTGAGAAAGTTCACTGAGGAGCAATATAC 416
QY 375 CCACTGGG 382
DB 417 CCACTGGG 424

RESULT 5
US-09-908-975-8960
; Sequence 8960, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simcha
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US/09/908,975
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8960
; LENGTH: 60
```

TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-908-975-8960

Query Match 10.6%; Score 60; DB 13; Length 60;  
Best Local Similarity 100.0%; Pred. No. 1,7e-08;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 TCAACCTGCTGTTATGTTAAAGCCCAAGACAGATCTATGAGGATTGGCACTTGG 253  
DB 1 TCAACCTGCTGTTATGTTAAAGCCCAAGACAGATCTATGAGGATTGGCACTTGG 60

## RESULT 6

US-09-908-975-1202  
Sequence 1202, Application US/09908975  
Publication No. US20030165843A1  
GENERAL INFORMATION:  
APPLICANT: SHOSHAN, Avi  
APPLICANT: WASSERMAN, Alon  
APPLICANT: MINTZ, Eli  
APPLICANT: MINTZ, Eli  
APPLICANT: FATIGLER, Simchon  
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE  
FILE REFERENCE: 36688-0005  
CURRENT FILING DATE: 2001-07-20  
PRIOR FILING DATE: 2001-05-02  
PRIOR APPLICATION NUMBER: US 60/287,724  
PRIOR FILING DATE: 2001-05-02  
PRIOR APPLICATION NUMBER: US 60/221,607  
NUMBER OF SEQ ID NOS: 2000-07-28  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 1202  
LENGTH: 65  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
US-09-908-975-1202

Query Match 7.2%; Score 41; DB 13; Length 65;  
Best Local Similarity 76.9%; Pred. No. 0.014;  
Matches 50; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 385 AGGGAGAAAGATCGTGTATGCTGAGGCCAAGATTCAATGATCCATCTGGGT 444  
DB 1 AGGGAGAAAGATCGTGTATGCTGAGGCCAAGATTCAATGATCCATCTGGGT 60

QY 445 CTTGG 449  
DB 61 CTTGG 65

## RESULT 7

US-09-918-995-30889  
Sequence 30889, Application US/09918995  
Publication No. US20030073623A1  
GENERAL INFORMATION:  
APPLICANT: HySeq, Inc.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
FILE REFERENCE: 20411-756  
CURRENT FILING DATE: 2001-07-30  
PRIOR FILING DATE: 1999-01-20  
PRIOR APPLICATION NUMBER: US/09/918,995  
NUMBER OF SEQ ID NOS: 38054  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 30889  
LENGTH: 484  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:

NAME/KEY: misc.feature  
LOCATION: (1)..(484)  
OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-30889

Query Match 6.9%; Score 39.4; DB 11; Length 484;  
Best Local Similarity 49.3%; Pred. No. 0.13;  
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 258 TACTGTGACCCGAGCTGATGCTCCAGAACTTGAAGAAAGCCATCCAAATTATGTA 317  
DB 231 TAGGTCTGTTCCAAAGCACTGTGCACAGAGCTTGAATTTCATTAGAGCCAAAC 290  
QY 318 CCAAACTGAGCAAGATGAGGCTGAGAAAGTTCACTGAGCCAGTGAATATACCCA 377  
DB 291 CCTGAGAAATACACAGGGGACGCTTCCAGTATGATGTTGGGAGAGAGGGCAG 350  
QY 378 CTGGAGAGGGAGAGAAATCAGCTGTGATGCTGAGCCCAAGATTATAGATGACAT 437  
DB 351 AGGGAGACAGGGGACAGGATTCAGCTTGTGTGGGTCTGAGGGGTTCTTACAGGGGTAG 410  
QY 438 CTTGGGTTCTGGCAGCAGCATTTGGGACTC 466  
DB 411 CCAAGATCTGGGAAACAGATCAGGACTC 439

## RESULT 8

US-10-232-484-5  
Sequence 5, Application US/10232484  
Publication No. US20030113847A1  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Yang, Xonghong  
APPLICANT: Drmanac, Radjle T  
TITLE OF INVENTION: 79ICIP2B2IV  
FILE REFERENCE: 79ICIP2B2IV  
CURRENT FILING DATE: 2002-08-30  
PRIOR FILING DATE: 2002-08-30  
PRIOR APPLICATION NUMBER: 09/695,783  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: 09/552,929  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5  
LENGTH: 1363  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (848)..(1090)  
OTHER INFORMATION:  
US-10-232-484-5

Query Match 6.9%; Score 39.4; DB 15; Length 1363;  
Best Local Similarity 49.3%; Pred. No. 0.24;  
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 258 TACTGTGACCCGAGCTGATGCTCCAGAACTTGAAGAAAGCCATCCAAATTATGTA 317  
DB 877 TAGGTCTGTTCCAAAGCACTGTGCACAGAGCTTGAATTTCATTAGAGCCAAAC 936  
QY 318 CCAAACTGAGCAAGATGAGGCTGAGAAAGTTCACTGAGCCAGTGAATATACCCA 377  
DB 937 CCTGAGAAATACACAGGGGACGCTTCCAGTATGATGTTGGGAGAGAGGGCAG 996  
QY 378 CTGGAGAGGGAGAGAAATCAGCTGTGATGCTGAGCCCAAGATTATAGATGACAT 437  
DB 997 AGGGAGACAGGGGACAGGATTCAGCTTGTGTGGGTCTGAGGGGTTCTTACAGGGGTAG 1056  
QY 438 CTTGGGTTCTGGCAGCAGCATTTGGGACTC 466  
DB 1057 CCAAGATCTGGGAAACAGATCAGGACTC 1085

```
RESULT 9
US-10-044-090-579
; Sequence 579, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 579
; LENGTH: 2319
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 023244.1
; NAME/KEY: unsure
; LOCATION: 169,173
; OTHER INFORMATION: a, c, g, or other
US-10-044-090-579

Query Match          6.9%; Score 39.4; DB 14; Length 2319;
Best Local Similarity 49.3%; Pred. No. 0.32;
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 258 TACTGTGGACCCAGACTGAGTGGCTCCAGAACCTAGAAAAGCCATCCAAATTATGTA 317
    |||||
DB 1560 TAGTGTCTGTTCCAAAGCACTGTGTACAGAGCTTAATTCATTAGAGGCACAAC 1619
    |||||

QY 318 CCAAAACCTGCAGCAAGATGGGCTGGAGAAAGTTCACTTGAGCCAGTGAATACCCCA 377
    |||||
DB 1620 CCTGAGAAATACAAAGGGGCGACGCTTCCAGTAGATGTGTTGGGAGAGAGGGCGAG 1679
    |||||

QY 378 CTGGAGAGGGGAGAGAAATCAGCTGTGATCTGAGCCAGCAAGAAATTCATAGATAGCCAT 437
    |||||
DB 1680 AGGGAGACAGGGGACAGATTCAGCTTGTGTGGTGGCTCTGAGGGTTCCTACAGGGGTAG 1739
    |||||

QY 438 CCTGGGCTTTGGCAGCAGCATTTGGGACTC 466
    |||||
DB 1740 CCAGATCTGGGAAACAGATCAGCGACTC 1768
    |||||

RESULT 10
US-09-814-353-20277
; Sequence 20277, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
```

```
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20277
; LENGTH: 3322
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-20277

Query Match          6.9%; Score 39.4; DB 13; Length 3322;
Best Local Similarity 49.3%; Pred. No. 0.4;
Matches 103; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 258 TACTGTGGACCCAGACTGAGTGGCTCCAGAACCTAGAAAAGCCATCCAAATTATGTA 317
    |||||
DB 1946 TAGTGTCTGTTCCAAAGCACTGTGTACAGAGCTTAATTCATTAGAGGCACAAC 2005
    |||||

QY 318 CCAAAACCTGCAGCAAGATGGGCTGGAGAAAGTTCACTTGAGCCAGTGAATACCCCA 377
    |||||
DB 2006 CCTGAGAAATACAAAGGGGCGACGCTTCCAGTAGATGTGTTGGGAGAGAGGGCGAG 2065
    |||||

QY 378 CTGGAGAGGGGAGAAATCAGCTGTGATCTGAGCCAGCAAGAAATTCATAGATAGCCAT 437
    |||||
DB 2066 AGGGAGACAGGGGACAGATTCAGCTTGTGTGGTGGCTCTGAGGGTTCCTACAGGGGTAG 2125
    |||||

QY 438 CCTGGGCTTTGGCAGCAGCATTTGGGACTC 466
    |||||
DB 2126 CCAGATCTGGGAAACAGATCAGCGACTC 2154
    |||||

RESULT 11
US-10-027-632-101927
; Sequence 101927, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101927
; LENGTH: 665
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-101927

Query Match          6.9%; Score 39; DB 13; Length 665;
Best Local Similarity 63.2%; Pred. No. 0.21;
Matches 60; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 9 ATATCTTAACAACAAACCACTGGAAGAAAAAATGGAATTCCTTATCTTCGATTGTT 68
    |||||
DB 389 AGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAATGAGATGAGCTGGCTTTGTT 448
    |||||

QY 69 CGATGATGTTACCTTTATCTCCCTGTGCTGGGA 103
    |||||
DB 449 TGTGGCTCACACCTGTAATTCAGCACTTTGGGA 483
    |||||
```

RESULT 12  
US-10-027-632-101927  
; Sequence 101927, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 101927  
; LENGTH: 665  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-101927

Query Match 6.9%; Score 39; DB 14; Length 665;  
Best Local Similarity 63.2%; Pred. No. 0.21;  
Matches 60; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Oy 9 ATTATCTTACAGAAACCACTGAGAAATGAAATTCCTTATCTTGCATTTT 68  
Db 389 AGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAATGGATGCTGGCTTTGTT 448  
Oy 69 CGGTGTGTTCACCTTTTATCCCTGTGCTGTGGA 103  
Db 445 TGGTGGCTCACACCTGTATTCAGCACTTTGGGA 483

RESULT 13  
US-10-412-277-3  
; Sequence 3, Application US/10412277  
; Publication No. US20030175791A1  
; GENERAL INFORMATION:  
; APPLICANT: GUEGLER, Karl et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CD001067D1V  
; CURRENT APPLICATION NUMBER: US/10/412,277  
; CURRENT FILING DATE: 2003-04-14  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 786431  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(786431)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-412-277-3

Query Match 6.5%; Score 36.6; DB 13; Length 786431;  
Best Local Similarity 50.3%; Pred. No. 64;

Matches 90; Conservative 0; Mismatches 89; Indels 0; Gaps 0;  
Oy 13 TCTTACAGAAAAACCACTGAGAAAAAATGAAATTCCTTATCTTGCATTTTGGCT 72  
Db 537863 TCACAAATTAAGAAAAAATAACAGAAAACTAAATTTGCTTAAAACTACTAATA 537922  
Oy 73 GGTGTTCACCTTTATCCCTGTGCTGTGGAAGCTATATGCAAGATGCACTCTTAAG 132  
Db 537923 AATTTTCAACTACTGTGCAAAACAAATCAAGAAAAATATAGTATAGTACAAAT 537982  
Oy 133 AGGACTTTGAAGAAATTAAGAAATAGCCAGCTGTGAGATGTTGCTTAAGCAAT 191  
Db 537983 ATTACAGATAAAAAGAAAAAGAAATATATAGTATTAATTAATTAATCAAT 538041

RESULT 14  
US-10-066-543-2072  
; Sequence 2072, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuguo  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Joseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Carter, Heather  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2072  
; LENGTH: 505  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-066-543-2072

Query Match 6.4%; Score 36.2; DB 15; Length 505;  
Best Local Similarity 60.8%; Pred. No. 1.3;  
Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Oy 7 AGATTATCTTACAGAAACCACTGAGAAAAAATGAAATTCCTTATCTTGCATTT 66  
Db 89 AATTAATTAATTAATGAAACAGATAGTAAATATATGAAATTTGTCATTCCTGCCGGG 148  
Oy 67 TTGGTGTGTTCACCTTTTATCCCTGTGCTGTGGA 103  
Db 149 TGTGTGGCTTAAGCCCTATATATCCAGCACTTTGGGA 185

RESULT 15  
US-09-764-991-9681  
; Sequence 9681, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antipodites  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; PRIOR application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9681  
; LENGTH: 1767  
; TYPE: DNA



ORGANISM: Homo sapiens  
US-09-764-891-9681

Query Match 6.4%; Score 36.2; DB 11; Length 1767;  
Best Local Similarity 60.8%; Pred. No. 2.7;  
Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy	7	AGATTATCTTACACAGAAACCACTGGAAAAAATGAATTCCTTATCTTCGCAATT	66
Db	481	AAATTAATTAATTAATGAAGAACAGATATAATTAATGAATTCCTTCGCGGG	540
Qy	67	TTGGTGCTTACCTTTATCCCTGCTCGGA	103
Db	541	TGTGTGCTTACGCTATAATCCAGCACTTGGGA	577

Search completed: December 22, 2003, 23:44:46  
Job time : 182.643 secs

**THIS PAGE BLANK (USPTO)**

QY 181 SRTVOYRTOGAVEAKVAGLALSLRSVASFISPHITGIOEDYVPIKPIACTITVEDAE 240  
 DB 181 SRTVOYRTOGAVEAKVAGLALSLRSVASFISPHITGIOEDYVPIKPIACTITVEDAE 240  
 QY 241 MMSRMASHGKIIVIQKMGAKTYPDTDSFNTVAEITGSKYPEQVVLVSGHLSMDVGGA 300  
 DB 241 MMSRMASHGKIIVIQKMGAKTYPDTDSFNTVAEITGSKYPEQVVLVSGHLSMDVGGA 300  
 QY 301 MDDGGGAFISWEALSLIKDLGLRPKRTLRVLTAEBOGGVGAFOYQOLHVNISNYSLV 360  
 DB 301 MDDGGGAFISWEALSLIKDLGLRPKRTLRVLTAEBOGGVGAFOYQOLHVNISNYSLV 360  
 QY 361 MMSDAGTFLPTGLOFTGSEKARAMEVWSLQPLNITQVLSHGEGTDINFWIQAGVPGA 420  
 DB 361 MMSDAGTFLPTGLOFTGSEKARAMEVWSLQPLNITQVLSHGEGTDINFWIQAGVPGA 420  
 QY 421 SLDDLYKYFFPHSHGDTMTVMDPKQNVAAVAVVSVVADMEMLPRS 472  
 DB 421 SLDDLYKYFFPHSHGDTMTVMDPKQNVAAVAVVSVVADMEMLPRS 472

RESULT 11

AAV87081  
 ID AAV87081 standard; Protein; 473 AA.

AAV87081;

09-MAY-2000 (first entry)

Human secreted protein sequence SEQ ID NO:120.

Human; secreted protein; diagnosis; cytostatic; immunosuppressive; anti-inflammatory; nootropic; neuroprotective; antiallergic; cancer; tumour; neurodegenerative disorder; developmental abnormality; allergy; foetal deficiency; blood disorder; immune system disorder; arthritis; autoimmune disease; hepatic disease; renal disease; inflammation; Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis; infection; AIDS; spinal cord injury; transplant rejection; diabetes; asthma; sepsis; acne; psoriasis; cardiovascular disorder; reproductive disorder; gastrointestinal disorder; respiratory disorder; metabolic disorder; food additive; preservative.

Homo sapiens.

WO200004140-A1.

27-JAN-2000.

14-JUL-1999; 99WO-US15849.

15-JUL-1998; 98US-0092921.

15-JUL-1998; 98US-0092922.

15-JUL-1998; 98US-0092956.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y, Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PB, Mucenski M, Endress GA, Soppet DR;

WPI: 2000-161128/14.

N-PSDB; AA298034.

New isolated human genes, useful for diagnosis and treatment of, e.g. cancers, neurological or blood disorders

Claim 11; Page 383-385; 494PP; English.

The polynucleotide sequences given in AA298017 to AA298108 encode the human secreted proteins given in AAV87064 to AAV87223. Human secreted protein can have activities based on the tissues and cells the genes are expressed in. Examples of activities include: cytostatic;

CC immunosuppressive; anti-inflammatory; nootropic; neuroprotective; and  
 CC antiallergic. The polynucleotides and their corresponding secreted  
 CC polypeptides are useful for preventing, treating or ameliorating medical  
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions  
 CC can be diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new  
 CC polynucleotides. Human secreted protein s and their polynucleotides can  
 CC be used for developing products for the diagnosis or treatment of cancer,  
 CC tumours, neurodegenerative disorders, developmental abnormalities and  
 CC foetal deficiencies, blood disorders, diseases of the immune system,  
 CC autoimmune diseases, hepatic and renal disease, inflammation,  
 CC allergies, Alzheimer's disease, behavioural disorders, schizophrenia,  
 CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,  
 CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,  
 CC cardiovascular disorders, reproductive disorders, gastrointestinal  
 CC disorders, respiratory disorders and metabolic disorders. The  
 CC proteins or polynucleotides can also be used as food additives or  
 CC preservatives. The proteins are also useful for identifying their  
 CC binding partners. AA298008 to AA298016 and AAV87063 are sequence used in  
 CC the exemplification of the present invention.

Sequence 473 AA:

Query Match 100.0%; Score 2417; DB 21; Length 473;

Best Local Similarity 100.0%; Pred. No. 6.8e-212;

Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAPFGVHLISLCSGKAICNGISKRTPEIKKESACGVAAIINLVYGAQ 60  
 DB 1 MKFLIFAPFGVHLISLCSGKAICNGISKRTPEIKKESACGVAAIINLVYGAQ 60  
 QY 61 NRSYERLALVDVTPRLSGSKNLEKAIQIMYONLQDGLKVLHPRIPIHMERGESBA 120  
 DB 61 NRSYERLALVDVTPRLSGSKNLEKAIQIMYONLQDGLKVLHPRIPIHMERGESBA 120  
 QY 121 VMLERIRIKIILIGSGSIGTPPEGITAEVIVVTSFDELQRRASBARGIIVVQPYNY 180  
 DB 121 VMLERIRIKIILIGSGSIGTPPEGITAEVIVVTSFDELQRRASBARGIIVVQPYNY 180  
 QY 181 SRTVOYRTOGAVEAKVAGLALSLRSVASFISPHITGIOEDYVPIKPIACTITVEDAE 240  
 DB 181 SRTVOYRTOGAVEAKVAGLALSLRSVASFISPHITGIOEDYVPIKPIACTITVEDAE 240  
 QY 241 MMSRMASHGKIIVIQKMGAKTYPDTDSFNTVAEITGSKYPEQVVLVSGHLSMDVGGA 300  
 DB 241 MMSRMASHGKIIVIQKMGAKTYPDTDSFNTVAEITGSKYPEQVVLVSGHLSMDVGGA 300  
 QY 301 MDDGGGAFISWEALSLIKDLGLRPKRTLRVLTAEBOGGVGAFOYQOLHVNISNYSLV 360  
 DB 301 MDDGGGAFISWEALSLIKDLGLRPKRTLRVLTAEBOGGVGAFOYQOLHVNISNYSLV 360  
 QY 361 MMSDAGTFLPTGLOFTGSEKARAMEVWSLQPLNITQVLSHGEGTDINFWIQAGVPGA 420  
 DB 361 MMSDAGTFLPTGLOFTGSEKARAMEVWSLQPLNITQVLSHGEGTDINFWIQAGVPGA 420  
 QY 421 SLDDLYKYFFPHSHGDTMTVMDPKQNVAAVAVVSVVADMEMLPRS 472  
 DB 421 SLDDLYKYFFPHSHGDTMTVMDPKQNVAAVAVVSVVADMEMLPRS 472

RESULT 12

AAW33604  
 ID AAW33604 standard; Protein; 472 AA.

AAW33604;

21-MAY-1998 (first entry)

Human secreted protein AM282 full-length sequence.

Secreted protein; AM282; cytokine; human.

Homo sapiens.

**THIS PAGE BLANK (USPTO)**

```

XX Key Location/Qualifiers
FH Peptide 1..24
FT /label= Sig_peptide
FT Protein 25..472
FT /label= Mat_protein
XX
XX WO9739030-A2.
XX
XX 23-OCT-1997.
XX
XX 16-APR-1997; 97WO-US06475.
XX
XX 13-JAN-1997; 97US-0783520.
XX
XX 18-APR-1996; 96US-0634325.
XX
XX (GENY ) GENETICS INST INC.
XX
XX Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA;
XX Spaulding V;
XX
XX MPI: 1997-526400/48.
XX
XX N-PSDB; AAV02286.
XX
XX New isolated secretory proteins AM340, AM282 and AK583 - possibly
XX have cytokine, cell proliferation/differentiation regulating,
XX immunomodulating activities, etc.
XX
XX Claim 16; Page 45-47; 59pp; English.
XX
XX This human secreted protein, designated AM282, is encoded by a
XX full-length cDNA clone (see AAV02286), deposited in ATCC 98026, that
XX was identified from a database search using an isolated partial
XX AM282 clone (see AAT97398). AM282 protein can be used in a claimed
XX method for preventing, treating or ameliorating a medical
XX condition. It may exhibit cytokine, cell proliferation (either
XX inducing or inhibiting) or cell differentiation (either inducing or
XX inhibiting) activity or may induce production of other cytokines in
XX certain cell populations. It may also exhibit e.g. immune
XX stimulating or suppressing activity, haematopoiesis regulating
XX activity, tissue growth activity, activin/inhibin activity,
XX chemotactic or chemokinetic activity, haemostatic or thrombolytic
XX activity, receptor/ligand activity, anti-inflammatory activity,
XX tumour inhibition activity, or other activities. No evidence of
XX any of these activities is given in the specification.
XX
XX Sequence 472 AA:
SQ
Query Match 99.7%; Score 2410; DB 18; Length 472;
Best Local Similarity 99.8%; Pred. No. 3e-211;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

DB 301 MDDGGAFISWEALSLIKDLGRLPKRTLRVLTAEBOGVGAFOYQLHKVNISNYSLV 360
QY 361 MESDAGTFLPTGLQFTSGSKRAIMEEYMSLQPLNTQVLSHREGDINFWIAGVPGA 420
DB 361 MESDAGTFLPTGLQFTSGSKRAIMEEYMSLQPLNTQVLSHREGDINFWIAGVPGA 420
QY 421 SLDDLYKFFPFHSHGDTMTVMDPKQNNVAAVAVSVVADMEELPRS 472
DB 421 SLDDLYKFFPFHSHGDTMTVMDPKQNNVAAVAVSVVADMEELPRS 472

RESULT 13
AAB10229
ID AAB10229 standard; Protein; 472 AA.
XX
XX AAB10229;
XX
XX 16-NOV-2000 (first entry)
XX
XX Human fetal kidney protein fragment AM282_1i.
XX
XX Secreted protein; cytostatic; immunostimulatory; antimicrobial;
XX antiviral; immunosuppressive; antiinflammatory; vulnertary; cytokine;
XX cell proliferation; differentiation; regulator; treatment; tumor;
XX autoimmune disease; inflammatory disorder; wound; microbial infection;
XX viral disease; graft versus host reaction suppression.
XX
XX Homo sapiens.
XX
XX WO200037630-A1.
XX
XX 29-JUN-2000.
XX
XX 22-DEC-1999; 99WO-US31005.
XX
XX 23-DEC-1998; 98US-0220876.
XX
XX (GENY ) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
XX Merberg D, Treacy M, Bowman MR;
XX
XX MPI: 2000-442661/38.
XX
XX N-PSDB; AAA40493.
XX
XX Secreted human proteins AS296-1i and AS34-1i, useful for treating
XX tumors, autoimmune diseases, inflammatory disorders, wounds, microbial
XX infections and viral diseases -
XX
XX Disclosure; Page 199-200; 293pp; English.
XX
XX This invention describes novel secreted human proteins (I) which have
XX cytostatic, immunostimulatory, antimicrobial, antiviral,
XX immunosuppressive, antiinflammatory and vulnertary activity and which act
XX as cytokine, cell proliferation or differentiation regulators. (I)
XX is useful for treating tumors, autoimmune diseases, inflammatory
XX disorders, wounds, microbial infections and viral diseases. (I) is also
XX useful for suppressing graft versus host reaction. AAB10226-B10288
XX represent the secreted proteins encoded by AAA40490-A40580 which are
XX described in the method of the invention.
XX
XX Sequence 472 AA:
SQ
Query Match 99.7%; Score 2410; DB 21; Length 472;
Best Local Similarity 99.8%; Pred. No. 3e-211;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

**THIS PAGE BLANK (USPTO)**

```

QY 261 LygThrTyPProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIleTyr 280
DB 786 AAGACCTACCCAGATGACTGATTCCTTCACACCTGTACAGAGACTGTGGAGCAAAATAT 845
QY 281 ProGluGlnValValIleValSerGlyHisIleuAspSerThrAspValGlyGlnGlyVal 300
DB 846 CCGAGACAGGTTGACTGCTGCTGACGTGACATCTGACAGCTGGAGTGTGGCAGGGTGC 905
QY 301 MetAspAspGlyGlyValAlaPheIleSerTrpGluAlaIleuSerIleuIleuAspLeu 320
DB 906 ATGGATGATGGCGGTGGAGCTTTATATCATGAGAGAGACTCTCACTTAATAAGACTT 965
QY 321 GlyLeuArgProIleAspGlnThrLeuArgLeuValIleuTrpThrAlaGluGlnGlnGly 340
DB 966 GGGGTGGCTCCAAAGAGAGACTCTCGGTGGTGTCTGACATGACAGAGAAAGAGTGA 1025
QY 341 ValGlyAlaPheGlnIleTyrGlnIleuHisIleValAsnIleSerAsnTyrSerIleuVal 360
DB 1026 GTTGATGCTTCCTCGATTTATCATCTTACACAGGTAATATTTCCAACTACAGTCTGTG 1085
QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLeu 380
DB 1086 ATGGAGTCTGACGACAGAACTCTTACCCAGCTGGCTGCAATTCACCTGGCAGTGAAG 1145
QY 381 AlaArgAlaIleMetGluGlnValIleuSerIleuGlnProLeuAsnIleThrGlnVal 400
DB 1146 GCGAGGGCGCATCAGAGAGAGAGTATAGCTGTGCGAGCCCTCCAAATATCAGCAGGTC 1205
QY 401 LeuSerHisGlyGlnGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyVal 420
DB 1206 CTGAGCCATGAGAGAGAGAGAGAGATCACTTTTGATCCAAAGCTGAGAGTCTGGAGCC 1265
QY 421 SerLeuLeuAspAspLeuTyrIleTyrPhePhePheHisIleSerHisGlyAspThrMet 440
DB 1266 AGTCTACTGATGATCTTATACAGATATTTCTTCTCCATCCTCCACGAGAGACACCATG 1325
QY 441 ThrValMetAspProIleGlnMetAsnValAlaAlaValIleTPAlaValIleSerTyr 460
DB 1326 ACTGTCAATGATCCAAAGAGAGATGAAAGTGTGCTGTGTGGCTGTGTGTTCTTAT 1385
QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
DB 1386 GTTGTTCAGACATGAGAGAAATGCTGCTAGTCC 1421

```

## RESULT 13

AAV02296 ID AAV02296 standard; DNA; 1778 BP.

AC AAV02296;

DT 21-MAY-1998 (first entry)

DB Human secreted protein AM282 full-length cDNA clone.

KM Secreted protein; AM282; cytokine; human; db.

OS Homo sapiens.

FT Key Location/Qualifiers

FT CDS 17..1435

FT sig\_peptide /+tag= a

FT mat\_peptide /+tag= b

FT /+tag= b

PD 23-OCT-1997.

PF 16-APR-1997; 97MO-US06475.

PR 13-JAN-1997; 97US-0783520.

```

PR 18-APR-1996; 96US-0634325.
XX
PA (GENY) GENETICS INST INC.
XX
PI Jacobs K, LaValle ER, McCoy JM, Werberg D, Racie LA;
PI Spaulding V;
XX
DR WPI, 1997-526400/48.
XX
DR P-PSDB; AAV03604.
XX
PT New isolated secretory proteins AM340, AM282 and AK583 - possibly
PT have cytokine, cell proliferation/differentiation regulating,
PT immunomodulating activities, etc.
XX
PS Claim 15; Page 44-45; 59pp; English.
XX
CC This cDNA clone encodes a protein (see M33604) designated AM282.
CC It was identified as "vfp5b10.r1 human EST 30142.5" (GenBank
CC accession No. R77830) in a database search using a partial AM282
CC clone (see T97398) obtained from a human foetal kidney cDNA
CC library using methods selective for cDNAs encoding secreted
CC proteins. AM282 is deposited in ATCC 98026 together with clones
CC AM340 (see T97397) and AK583 (see V02297), which are also claimed.
CC AM282 protein can be used in a claimed method for preventing,
CC treating or ameliorating a medical condition. It may exhibit
CC cytokine, cell proliferation (either inducing or inhibiting) or
CC cell differentiation (either inducing or inhibiting) activity or
CC may induce production of other cytokines in certain cell
CC populations. It may also exhibit e.g. immune stimulating or
CC suppressing activity, haematopoiesis regulating activity, tissue
CC growth activity, activin/inhibin activity, chemotactic or
CC chemokinetic activity, haemostatic or thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, tumour
CC inhibition activity, or other activities. No evidence of any of
CC these activities is given in the specification.
XX
SQ - Sequence 1778 BP; 514 A; 386 C; 422 G; 456 T; 0 other;

```

## Alignment Scores:

Pred. No.: 1,02e-235 Length: 1778  
 Score: 2410.00 Matches: 471  
 Percent Similarity: 99.79% Conservative: 0  
 Best Local Similarity: 99.79% Mismatches: 1  
 Query Match: 99.71% Gaps: 0  
 DB: 18

US-09-745-763-36 (1-472) x AAV02296 (1-1778)

```

QY 1 MetIysPheLeuIlePheAlaPhePheGlyValHisIleuSerLeuCySerGly 20
DB 17 ATGAATTCCTTATCTTCGCAATTTTCGGTGGTGTTCACCTTTATCCCTGTGCTGGG 76
QY 21 LysAlaIleCysIysAsnGlyIleSerIysArgThrPheGluGluIleuGluGlnIle 40
DB 77 AAGACTATATGCAAGATGAGATGCACTCTTAAGAGACTTTTAAGAAATTAAGAAATA 136
QY 41 AlaserCysGlyAspValAlaIleValIleAsnLeuAlaValIleTyrGlyValGln 60
DB 137 GCGAGCTGTGAGATGTTGCTTAAGCAATCATCACTTACTGTTATGTAAGCCAG 196
QY 61 AsnArgSerTyrGluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80
DB 197 AACGATCTTATGAGCATTTGAGCACTTGTGATATCTGTGAGACCAAGCTGAGTGC 256
QY 81 SerIysAsnLeuGluIysAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100
DB 257 TCCAAAGACTTAAGAAAGCAATCCAAATTTATGTAACCAAACTGCGCAAGAGGCTG 316
QY 101 GluIysValHisLeuGluProValArgIleProHisIleTrpGluArgGlyGluSerAla 120
DB 317 GAGAAAGTCACTCGAGCGAGAGATATCCCACTGCGAGAGAGGAGAAAGATCAGCT 376
QY 121 ValMetLeuGluProArgIleHisIleValIleLeuValIleuGlySerSerIleGly 140

```

**THIS PAGE BLANK (USPTO)**



```

Db      377  GTATGCTGAGCCAGAAATTCATAGATGCCCTCTGGCTTGCACAGCATGGG 436
Qy      141  ThrProGluGlyIleThrAlaGluValIleuValThrSerPheAspGluLeuGln 160
Db      437  ACTCTCCAGAAAGCATTAACGCAAGATTCTGGTGTGACCTCTTTCGATGAACTGCAG 496
Qy      161  ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180
Db      497  AGAAGGCGCTCAGAAAGCAGAGAGAGATTGTTGTTATACCACTTATCATCACTAC 556
Qy      181  SerArgThrValGlnTyrArgThrGlnGlyValAlaGluAlaAlaLysValGlyAlaLeu 200
Db      557  TCAAGACGCTGCAATATCCAAACGAGGGCGGTGGAACCTGCACAGGGGGGCTTTG 616
Qy      201  AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHISThrGlyIleGln 220
Db      617  GCATCTCTCATTTGATCCGTCGCTCTTCTTCATCTACAGTCTCTCACAGGATTTGAG 676
Qy      221  GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db      677  GAATACCAAGATGGCGGCCCAAGATTCACACAGCCTGATTTACGGTGAAGATGCAGAA 736
Qy      241  MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db      737  ATGATGTCAAGATGGCTTCTCTCATGGATCAAAATTTCTCATGAGCTAAAGATGGGGCA 796
Qy      261  LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
Db      797  AAGACTTACCCAGATCTGATTTCTTCTTACACACTGTACAGAGATCACTGGAGCAAAATAT 856
Qy      281  ProGluGlnValIleuValSerGlyHisLeuAspSerTyrAspValGlyGlnGlyAla 300
Db      857  CCAGAACAGATGTGACTGCTGACTGACATCTGACAGCTGGAGTGTGGGAGGATGCC 916
Qy      301  MetAspAspGlyGlyGlyValAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320
Db      917  ATGAGATGAGGGCGGTGAGCTTTATATATCATGGGAAGCACTTCTCATTAAGATCTT 976
Qy      321  GlyLeuArgProLysArgThrLeuArgLeuValIleuTyrThrAlaGluGlnGlnGly 340
Db      977  GGGCTCCGTCCAAGAGCACTCTGCGGCTGGTCTCTGACCTGCAGAGAAACAGGTGGA 1036
Qy      341  ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
Db      1037  GTTGGCTCTTCAGATTATCAGTTACACAAAGTAAATTTCCAACTACAGTCTGCTG 1096
Qy      361  MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLys 380
Db      1097  ATGGAGTCTGAGCGAGGAACCTTCTTACCCACTGGGCTCAATTCTGCGCAGTGAAG 1156
Qy      381  AlaArgAlaIleMetGluGlnValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db      1157  GCCAAGGCGCATATGAGAGAGGTTATGACCTGCTGCAACCTCAATATCACTCGAGTC 1216
Qy      401  LeuSerHisGlyGlyGlyThrAspIleAsnPheTyrIleGlnAlaGlyValProGlyAla 420
Db      1217  CTGAGCATGAGAGAGGAGCAGACATCACTTTGGATCCAACTGAGATGCTGAGGCC 1276
Qy      421  SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440
Db      1277  AGCTACTTGAAGACTTATACAAATATTTCTTCTTCATCACTCCCAAGAGACACCATG 1336
Qy      441  ThrValMetAspProLysGlnMetAsnValAlaAlaValTyrAlaValAspTyr 460
Db      1337  ACTGTCATGATCCAAAGCAGATGAAATGTGCTGCTGCTGTTGGCTGTGTTCTTAT 1396
Qy      461  ValValAlaAspMetGluGlnMetLeuProArgSer 472
Db      1397  GTTGTTCAGACATGAGAAATGCTGCTAGTCTC 1432

```

RESULT 14  
AAK94491

```

ID      AAK94491 standard; cDNA; 1860 BP.
XX
AC      AAK94491;
XX
DT      06-NOV-2001 (first entry)
XX
DE      Human full-length cDNA, SEQ ID NO: 3328.
XX
KW      Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS      Homo sapiens.
XX
PN      EP1130094-A2.
XX
PD      05-SEP-2001.
XX
PF      07-JUL-2000; 2000EP-0114089.
XX
PR      08-JUL-1999; 99JP-0194486.
PR      11-JAN-2000; 2000JP-0118774.
PR      02-MAY-2000; 2000JP-0183765.
XX
PA      (HELI-) HELIX RES INST.
XX
PI      Ota T, Nishikawa T, Isegai T, Hayashi K, Ishii S, Kawai Y;
PI      Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR      WPI, 2001-524255/58.
XX
PT      P-PSDB; AAM93559.
XX
PT      830 Primers useful for synthesizing full length cDNA clones and their
PT      use in genetic manipulation -
XX
PS      Claim 8, SEQ ID NO 3328; 1380bp + sequence listing; English.
XX
CC      The invention relates to primers for synthesizing full length cDNA
CC      clones. 830 cDNA molecules encoding a human protein have been
CC      isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC      molecules have been determined. Primers for synthesizing the full length
CC      cDNA are useful for clarifying the function of the protein encoded by
CC      the cDNA. The full length clones were obtained by construction of full
CC      length enriched cDNA libraries that were synthesised by the oligo-capping
CC      method. The primers enable the production of the full length cDNA easily
CC      without any special methods. The present sequence is a full length
CC      human cDNA of the invention.
CC
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ      Sequence 1860 BP; 515 A; 423 C; 451 G; 471 T; 0 other;

Alignment Scores:
Pred. No.: 7,146-235 Length: 1860
Score: 2402.00 Matches: 469
Percent Similarity: 99.79% Conservative: 2
Best Local Similarity: 99.36% Mismatches: 1
Query Match: 99.38% Indels: 0
Gaps: 0

US-09-745-763-36 (1-472) x AAK94491 (1-1860)
Qy      1  MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCysSerGly 20
Db      129  ATGAATATCTTATCTTGGCAATTTTGGTGGTGTACCTTTATATCCCTGTGCTGGG 188
Qy      21  LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluLe 40
Db      189  AAAGCTATATCAAGAAAGATGCTCTTAAGAGCACTTTTGAAGAAATGAAGAAATTA 248
Qy      41  AlaSerCysGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60
Db      249  GCCAGCTGTGAGAGATGTTGCTAAAGCAATCAATCAACCTAGCTTTATGTAAGCCAG 308
Qy      61  AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:34:19 ; Search time 80.2473 Seconds  
(without alignments)  
933.601 Million cell updates/sec

Title: US-09-745-763-36  
Perfect score: 2417  
Sequence: 1 MKELFAFGVHLLSLICSG.....AVNAVSVVADMEMLPRS 472

Scoring table: BLASTSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq.19Jun03:\*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*
- 24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2417	100.0	472	19	AAW85456
2	2417	100.0	472	21	AAW87254
3	2417	100.0	472	21	AAW58879
4	2417	100.0	472	22	AAW60658
5	2417	100.0	472	23	ABP61800
6	2417	100.0	472	23	ABG33880
7	2417	100.0	472	24	ABR47892
8	2417	100.0	472	24	ABR48106
9	2417	100.0	472	24	ABR00152

10	2417	100.0	472	24	ABR00274	Human gene 142 enc
11	2417	100.0	473	21	AAW87081	Human secreted pro
12	2410	99.7	472	18	AAW33604	Human secreted pro
13	2410	99.7	472	21	AAW10329	Human fetal kidney
14	2402	99.4	472	22	AAW93559	Human polypeptide
15	2378.5	98.4	461	22	AAW25762	Human protein sequ
16	2340	96.8	474	22	ABG09405	Novel human diagno
17	899	37.2	178	24	ABG48107	Human secreted pro
18	899	37.2	178	24	ABR00275	Human gene 142 enc
19	460	19.0	149	18	AAW27642	Secreted protein A
20	460	19.0	149	18	AAW31633	Human secreted pro
21	460	19.0	149	18	AAW44078	Human secreted pro
22	397	16.4	90	22	AAW84050	Human immune/haema
23	351	14.5	71	20	AAW11865	Human 5' EST secre
24	229.5	9.5	496	20	AAW89614	Aspergillus oryzae
25	229.5	9.5	496	20	AAW89597	Aspergillus oryzae
26	229.5	9.5	496	20	AAW89586	Aspergillus oryzae
27	226.5	9.4	481	21	AAW58020	Human peptidase NA
28	188	7.8	740	21	AAW58878	Human EST encoded
29	188	7.8	740	22	AAW24008	Human EST encoded
30	188	7.8	740	22	AAW88354	Human membrane or
31	173	7.2	750	21	AAW92644	Mutant human prost
32	169	7.0	333	17	AAW9416	Aminopeptidase pre
33	169	7.0	752	21	AAW92623	Murine prostate sp
34	169	7.0	761	21	AAW92650	Mutant murine pros
35	168	7.0	750	21	AAW92628	Mutant human prost
36	168	7.0	751	18	AAW31524	Prostate-specific
37	167.5	6.9	600	21	AAW21943	Arabidopsis thalia
38	167.5	6.9	610	21	AAW21942	Arabidopsis thalia
39	166	6.9	693	21	AAW92647	Mutant human PSM a
40	166	6.9	766	21	AAW92659	Mutant murine pros
41	165	6.8	33	21	AAW87185	Human secreted pro
42	165	6.8	33	22	AAW06125	Human gene 18 enco
43	165	6.8	33	22	AAW06162	Human gene 18 enco
44	165	6.8	33	23	ABG33947	Human secreted pro
45	165	6.8	33	23	ABG33985	Human secreted pro

## ALIGNMENTS

RESULT 1

AAW85456 standard; Protein; 472 AA.

AAW85456;

25-FEB-1999 (first entry)

Secreted protein encoded by clone bu45\_2.

Secreted protein; nutritional activity; immune stimulating; vaccine;

suppressing activity; haematopoiesis regulating activity;

tissue growth activity; activin; inhibin activity; chemotaxis;

chemokinetic activity; haemostasis; thrombolytic activity; receptor;

ligand; anti-inflammatory; cadherin; tumour invasion suppressor;

tumour inhibition; gene therapy.

XX Homo sapiens.

XX OS

XX WO9842739-A2.

XX PN

XX 01-OCT-1998.

XX PD

XX 20-MAR-1998; 98WO-US05653.

XX PF

XX 19-MAR-1998; 98US-0044466.

XX PR

XX 21-MAR-1997; 97US-0822167.

XX PR

XX (GENY) GENETICS INST INC.

XX PA

XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Werberg D;

XX PI Racie IA, Spaulding V, Treacy M;





KW Human; secreted protein; proliferative disorder; cancer; tumour; asthma;  
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;  
 KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;  
 KW inflammation; neurological disorder; Alzheimer's disease; food additive;  
 KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;  
 KW pregnancy-related disorder; endocrine disorder; infection; wound healing;  
 KW cell culture; chemotaxis; vulnery; binding partner identification;  
 KW gene therapy.  
 KW  
 OS Homo sapiens.  
 XX  
 XX  
 XX Key Location/Qualifiers  
 FT Peptide 1..24  
 FT /label= Signal\_peptide  
 FT 25..472  
 FT /label= Mature\_human\_secreted\_protein  
 XX  
 XX MO200151504-A1.  
 XX  
 XX 19-JUL-2001.  
 XX  
 XX 12-JAN-2001; 2001WO-US00911.  
 XX  
 XX 13-JAN-2000; 2000US-0482273.  
 XX  
 XX (HUMAN-) HUMAN GENOME SCI INC.  
 XX  
 XX Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y,  
 PI Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR,  
 PI Endress GA, Muscenski M, Ebner R;  
 XX  
 XX WPI: 2001-425865/45.  
 DR N-BSDB; AAD11647.  
 XX  
 XX Isolated nucleic acid molecule encoding a human secreted protein is  
 PT used in preventing, treating or ameliorating a medical condition -  
 XX  
 PS Claim 11; Page 750-752; 864pp; English.  
 XX  
 XX AAD1163C-AAD11721 represent cDNAs corresponding to 71 human secreted  
 CC protein genes, and AAE06041-AAE06132 represent the proteins they encode.  
 CC AAE06133-AAE06205 represent human secreted protein fragments.  
 CC The secreted proteins and their genes are useful for preventing, treating  
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
 CC Pathological conditions can be diagnosed by determining the amount of the  
 CC new protein in a sample or by determining the presence of mutations in  
 CC the new genes. Specific uses are described for each of the 71 genes,  
 CC based on the tissues in which they are most highly expressed, and include  
 CC developing products for the diagnosis or treatment of proliferative  
 CC disorders; cancer; tumours; foetal and developmental abnormalities;  
 CC haematopoietic disorders; diseases of the immune system; AIDS; autoimmune  
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,  
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,  
 CC pregnancy-related disorders, endocrine disorders, and infections. The  
 CC proteins can also be used to aid wound healing and epithelial cell  
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
 CC before transplantation, for supporting cell culture of primary tissues,  
 CC to regenerate tissues, to identify their cognate ligands or binding  
 CC partners, and in chemotaxis, and can be used as a food additive or  
 CC preservative to modify storage properties. Antibodies specific for a  
 CC protein of the invention can be used in alleviating symptoms associated  
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).  
 CC The present sequence represents a human secreted protein of  
 CC the invention.  
 CC  
 XX Sequence 472 AA;  
 XX  
 XX 50

Query Match 100.0%; Score 2417; DB 22; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 6, 8e-212;  
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKFLIPAFGGVHLLSLCSGKAIKNGISKRTFEEIKBEIASGCDVAKIINLAVYGAQ 60  
 DB 1 MKFLIPAFGGVHLLSLCSGKAIKNGISKRTFEEIKBEIASGCDVAKIINLAVYGAQ 60  
 QY 61 NRSYERLALVDYTPGLSSGKNEKAIQIMYQNLQODGKEKVLPEPRI PMMERGESA 120  
 DB 61 NRSYERLALVDYTPGLSSGKNEKAIQIMYQNLQODGKEKVLPEPRI PMMERGESA 120  
 QY 121 VMEPRRIHKIALILGSSIGTPPEGITAEVLVVSFPDLQRASEARGKIVVNYQPIYNY 180  
 DB 121 VMEPRRIHKIALILGSSIGTPPEGITAEVLVVSFPDLQRASEARGKIVVNYQPIYNY 180  
 QY 181 SRTVQYRTQGAVERAKVAGALASLIRSVASFISYPHGTIOEQGVKIPACTIVDAE 240  
 DB 181 SRTVQYRTQGAVERAKVAGALASLIRSVASFISYPHGTIOEQGVKIPACTIVDAE 240  
 QY 241 MMSRMASHGIKIVIQMGAKATYPTDSFNTVAETISGSKYPEQVVLVSGHDSMDVGGA 300  
 DB 241 MMSRMASHGIKIVIQMGAKATYPTDSFNTVAETISGSKYPEQVVLVSGHDSMDVGGA 300  
 QY 301 MDDGGCAPIISWEALSLIKDLGRPKRTLRLVMTAEBOGVGAVQYQLHKVNTSNYSLV 360  
 DB 301 MDDGGCAPIISWEALSLIKDLGRPKRTLRLVMTAEBOGVGAVQYQLHKVNTSNYSLV 360  
 QY 361 MESDAGTFPLPTGLQFTSGSEKARAIMBEVMSILQPLNITQVLSHGEGTDINFIQAGVGA 420  
 DB 361 MESDAGTFPLPTGLQFTSGSEKARAIMBEVMSILQPLNITQVLSHGEGTDINFIQAGVGA 420  
 QY 421 SLDDLVKYPFFHHSHDPTVMDPKMNVAAVAIVSVVVAAMEELPRS 472  
 DB 421 SLDDLVKYPFFHHSHDPTVMDPKMNVAAVAIVSVVVAAMEELPRS 472  
 RESULT 5  
 ABE61800  
 ID ABE61800 standard; Protein; 472 AA.  
 XX  
 XX ABE61800;  
 AC  
 XX  
 DT 04-OCT-2002 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 154.  
 XX  
 KW Human; cytosolic; antirheumatic; antiarthritic; vulnery; analgesic;  
 KW antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;  
 KW neuroprotective; noctropic; osteopathic; haemostatic; vasotropic;  
 KW antitumor; fungicide; antidiabetic; antiaesthetic; antiallergic;  
 KW immunostimulant; antiparasitic; secreted protein; transmembrane protein;  
 KW cytokine; cell proliferation; cell differentiation; autoimmune disease;  
 KW stem cell; growth factor; nervous system disease; neuropathy;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;  
 KW osteoporosis; severe combined immunodeficiency; SCID; infection;  
 KW multiple sclerosis; rheumatoid arthritis; gene therapy.  
 KW  
 OS Homo sapiens.  
 XX  
 XX US2002065394-A1.  
 XX  
 XX 30-MAY-2002.  
 XX  
 XX 22-DEC-2000; 2000US-0745763.  
 XX  
 XX 18-MAR-1998; 98US-0040963.  
 XX  
 XX (JACOBS K.  
 PA (MCCOY J M.  
 PA (LAVALLIE E R.  
 PA (COLLINS-RACIE L A.  
 PA (EVANS C.  
 XX

PA (MERB/) MERBERG D.  
 PA (TREAC/) TREACY M.  
 PA (SPAUV/) SPAULDING V.  
 XX  
 PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Spaulding V;  
 XX  
 DR WPI: 2002-582343/62.  
 DR N-PSDB: ABQ92016.  
 XX  
 PT Novel secreted or transmembrane protein and polynucleotide encoding the  
 PT protein, useful for diagnosis and treatment of neurological disorders,  
 PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis  
 PT  
 XX  
 PS Claim 51; Page 114-115; 264pp; English.  
 XX  
 CC The invention relates to human secreted or transmembrane protein (I),  
 CC their fragments and is encoded by specific complementary deoxyribonucleic  
 CC acid (CDNA) inserts (II), where the protein is substantially free from  
 CC other mammalian proteins. (I) are useful for preventing, treating or  
 CC ameliorating a medical condition, especially immunological treatment or  
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,  
 CC cytokine, cell proliferation, cell differentiation, antiinflammatory,  
 CC stem cell growth factor activity and activin or inhibin-related  
 CC activities. (I) can be used to manipulate stem cells in culture to give  
 CC rise to neuroepithelial cells that can be used to augment or replace  
 CC cells damaged by illness, autoimmune disease, accidental damage or  
 CC genetic disorders. (I) induces the proliferation of neural cells and  
 CC regeneration of nerve and brain tissue and is useful for the treatment of  
 CC central and peripheral nervous system diseases and neuropathies, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic  
 CC activity, regulation of haematopoiesis and is useful for treating myeloid  
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopaenia  
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,  
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or  
 CC periodontal disease. (I) is also useful for gut protection or  
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury  
 CC in various tissues, various immune deficiencies and disorders including  
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,  
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,  
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,  
 CC such as asthma or other respiratory problems. (II) is useful to express  
 CC recombinant protein, as markers for tissues in which the corresponding  
 CC protein is preferentially expressed and in gene therapy. The present  
 CC sequence is that of a polypeptide of the invention.  
 CC  
 XX  
 SQ Sequence 472 AA;  
 Query Match 100.0%; Score 2417; DB 23; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-212;  
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKPLIFAFGGVHLISLCSGKAIQNGISKRTPEIKKEIASCGVAAKINILNAYGAAQ 60  
 DB 1 MKPLIFAFGGVHLISLCSGKAIQNGISKRTPEIKKEIASCGVAAKINILNAYGAAQ 60  
 QY 61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYONLODGLKLVHLEPVRIPIHMERGESGA 120  
 DB 61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYONLODGLKLVHLEPVRIPIHMERGESGA 120  
 QY 121 VMLBPRIRIKIILIGSSIGTPPEGITLVEVLVYTSFDELORRASEARGIIVVNOPTINY 180  
 DB 121 VMLBPRIRIKIILIGSSIGTPPEGITLVEVLVYTSFDELORRASEARGIIVVNOPTINY 180  
 QY 181 SRTVYRTQGAWEAKVAGALASLRVASFISYSHGTGOEODVPEKIPACTIVVEAE 240  
 DB 181 SRTVYRTQGAWEAKVAGALASLRVASFISYSHGTGOEODVPEKIPACTIVVEAE 240  
 QY 241 MMSRMASHGIKIVIQLKMGAKTYPTDTSFNTVAEITGSKYPEQVVLVSGHLSMDVQGA 300  
 DB 241 MMSRMASHGIKIVIQLKMGAKTYPTDTSFNTVAEITGSKYPEQVVLVSGHLSMDVQGA 300

DB 241 MMSRMASHGIKIVIQLKMGAKTYPTDTSFNTVAEITGSKYPEQVVLVSGHLSMDVQGA 300  
 QY 301 MDDGGGAFISWEALSLIKDGLRPKRTLRLVLTAEBOGGVGAFOYYQLHKVINSNLSV 360  
 DB 301 MDDGGGAFISWEALSLIKDGLRPKRTLRLVLTAEBOGGVGAFOYYQLHKVINSNLSV 360  
 QY 361 MESDAGTFLPTGLQFTGSEKARAIMEEVMSLQPLNTQVLSHGEQDINFWIQAGVPGA 420  
 DB 361 MESDAGTFLPTGLQFTGSEKARAIMEEVMSLQPLNTQVLSHGEQDINFWIQAGVPGA 420  
 QY 421 SLDDLKYRPFPHSHGDTMTWMDPKOMNTAAAVMAVSVVADMEMLPRS 472  
 DB 421 SLDDLKYRPFPHSHGDTMTWMDPKOMNTAAAVMAVSVVADMEMLPRS 472  
 RESULT 6  
 ABG33880  
 ID ABG33880 standard; Protein: 472 AA.  
 XX  
 AC ABG33880;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human secreted protein encoded by gene 18 #1.  
 XX  
 KW Human; secreted protein; gene therapy; immunosuppressive;  
 KW antiarthritis; antirheumatic; antiproliferative; cyostatic; cardiant;  
 KW vasotropic; cerebroprotective; motropic; neuroprotective; antibacterial;  
 KW vinca; fungicide; ophthalmological; autoimmune disease; neoplasm;  
 KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;  
 KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;  
 KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
 KW ocular disorder; corneal infection; wound healing; skin aging;  
 KW epithelial cell proliferation; food additive.  
 KW  
 XX  
 OS Homo sapiens.  
 OS  
 XX  
 PN MO200226931-A2.  
 PD 04-APR-2002.  
 PD  
 PP 24-SEP-2001; 2001WO-US29871.  
 PP  
 XX 25-SEP-2000; 2000US-234925P.  
 PR 12-JAN-2001; 2001WO-US00911.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y;  
 PI Lafleur DW, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;  
 PI Endress GA, Mucenski M, Ebner R;  
 XX  
 DR WPI: 2002-362489/39.  
 DR N-PSDB: ABK69743.  
 XX  
 PT Novel 71 isolated secreted polypeptides and polynucleotides encoding  
 PT the polypeptides, useful for treating Huntington's disease, sepsis,  
 PT meningitis, thrombocytopaenia, haemolytic anaemia, rheumatoid arthritis,  
 PT asthma  
 PT  
 PS Claim 11; Page 1231-1232; 1478pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (or its  
 CC fragment, homologue complement or allelic variant) encoding a human  
 CC secreted protein (and its fragment, domain, epitope, variant, secreted  
 CC form and species variant). Also included are a recombinant vector  
 CC comprising the nucleic acid, a recombinant host cell comprising the  
 CC vector, an antibody against the secreted protein, a recombinant host cell  
 CC that expresses the secreted protein and a method of identifying a binding  
 CC partner of the secreted protein. The nucleic acid and protein are used to  
 CC prevent, diagnose, treat or ameliorate a medical condition in e.g.  
 CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep  
 CC for example autoimmune diseases e.g. rheumatoid arthritis,

hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. Many other diseases and disorders are listed in the specification. The polypeptides can also be used to aid wound healing in epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. The present sequence represents a novel human secreted protein of the invention.

Sequence 472 AA;

Query Match 100.0%; Score 2417; DB 23; Length 472;  
Best Local Similarity 100.0%; Pred. No. 6.8e-212;  
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MKFLIFAFPGVHLLSLCSGKAIQIMYQNLQODGLEKVLHPYRIPMERGESB  
1 NRSYERLALVDVTPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPYRIPMERGESB  
61 NRSYERLALVDVTPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPYRIPMERGESB  
61 NRSYERLALVDVTPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPYRIPMERGESB  
121 VMLERIRHKAIIIGLSSIGTPPEGITAEVLYVTSFDELQRRASRAGKIIVVNOPTINY  
121 VMLERIRHKAIIIGLSSIGTPPEGITAEVLYVTSFDELQRRASRAGKIIVVNOPTINY  
181 SRTVOYRTOGAVEAKVGLASLRSVASFISYSPHTGISOEDQVPIPTACTIVEDAE  
181 SRTVOYRTOGAVEAKVGLASLRSVASFISYSPHTGISOEDQVPIPTACTIVEDAE  
241 MESDAGTFLPTGLQFTSEKARAIMBEVMSLLQPLNITQVLSHGEGTDINFWIQGVGA  
241 MESDAGTFLPTGLQFTSEKARAIMBEVMSLLQPLNITQVLSHGEGTDINFWIQGVGA  
301 MDDGGAFISWEALSLIKDLGLRPKRTLRVLTAEBOGGVGAFQYQLHKNVINSYSLV  
301 MDDGGAFISWEALSLIKDLGLRPKRTLRVLTAEBOGGVGAFQYQLHKNVINSYSLV  
361 MESDAGTFLPTGLQFTSEKARAIMBEVMSLLQPLNITQVLSHGEGTDINFWIQGVGA  
361 MESDAGTFLPTGLQFTSEKARAIMBEVMSLLQPLNITQVLSHGEGTDINFWIQGVGA  
421 SLDDLYKYFFPHSHGDTMTVMDPKQNVAAVAVVSYVADMEMLPRS 472  
421 SLDDLYKYFFPHSHGDTMTVMDPKQNVAAVAVVSYVADMEMLPRS 472

RESULT 7

ABR47892 standard; Protein; 472 AA.

ABR47892;

12-JUN-2003 (first entry)

Human secreted protein, SEQ ID 783.

Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;

vulnerary; antiinflammatory; neuroprotective;

antiparkinsonian; gene therapy; human; cardiovascular disorder.

Homo sapiens.

MO200295010-12.

28-NOV-2002.

19-MAR-2002; 2002WO-US09785.

21-MAR-2001; 2001US-277340P.  
19-JUN-2001; 2001US-306171P.  
13-NOV-2001; 2001US-331287P.  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Ruben SM;  
WPI; 2003-129429/12.

Novel human secreted proteins, useful for detecting, preventing, diagnosing, prognosticating, treating and/or ameliorating cardiovascular disorders such as arrhythmia.

Claim 13; SEQ ID 783; 1881bp; English.

The present invention relates to novel human secreted proteins (ABR47633-ABR48145) and their coding sequences (ACC50344-ACC50856). The proteins and their coding sequences are useful for the preparation of a diagnostic or pharmaceutical composition for diagnosing or treating a cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary arteriosclerosis and myocardial ischaemia), neural disorders, immune system disorders, muscular disorders, reproductive disorders, gastrointestinal disorders, pulmonary disorders, renal disorders, proliferative disorders and/or cancerous diseases and conditions, for wound healing and epithelial cell proliferation, to treat inflammation or infection, for creating thrombosis and arteriosclerosis, for treating or preventing neural damage which occurs in neuronal disorders or neurodegenerative conditions such as Alzheimer's disease and Parkinson's disease, to enhance bone and periodontal regeneration and aid in tissue transplants or bone grafts, to prevent skin aging or hair loss, to stimulate growth and differentiation with other cytokines, to maintain organs before transplantation or for supporting cell culture of primary tissues, to increase or decrease differentiation or proliferation of embryonic stem cells, or to modulate mammalian characteristics or metabolism.

Note: The sequence data for this patent was published in electronic format and is available from WIPO at [ftp://wipo.int/pub/published\\_pcl\\_sequences](http://wipo.int/pub/published_pcl_sequences).

Sequence 472 AA;

Query Match 100.0%; Score 2417; DB 24; Length 472;  
Best Local Similarity 100.0%; Pred. No. 6.8e-212;  
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MKFLIFAFPGVHLLSLCSGKAIQIMYQNLQODGLEKVLHPYRIPMERGESB  
1 MKFLIFAFPGVHLLSLCSGKAIQIMYQNLQODGLEKVLHPYRIPMERGESB  
61 NRSYERLALVDVTPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPYRIPMERGESB  
61 NRSYERLALVDVTPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPYRIPMERGESB  
121 VMLERIRHKAIIIGLSSIGTPPEGITAEVLYVTSFDELQRRASRAGKIIVVNOPTINY  
121 VMLERIRHKAIIIGLSSIGTPPEGITAEVLYVTSFDELQRRASRAGKIIVVNOPTINY  
181 SRTVOYRTOGAVEAKVGLASLRSVASFISYSPHTGISOEDQVPIPTACTIVEDAE  
181 SRTVOYRTOGAVEAKVGLASLRSVASFISYSPHTGISOEDQVPIPTACTIVEDAE  
241 MESDAGTFLPTGLQFTSEKARAIMBEVMSLLQPLNITQVLSHGEGTDINFWIQGVGA  
241 MESDAGTFLPTGLQFTSEKARAIMBEVMSLLQPLNITQVLSHGEGTDINFWIQGVGA  
301 MDDGGAFISWEALSLIKDLGLRPKRTLRVLTAEBOGGVGAFQYQLHKNVINSYSLV  
301 MDDGGAFISWEALSLIKDLGLRPKRTLRVLTAEBOGGVGAFQYQLHKNVINSYSLV  
361 MESDAGTFLPTGLQFTSEKARAIMBEVMSLLQPLNITQVLSHGEGTDINFWIQGVGA  
361 MESDAGTFLPTGLQFTSEKARAIMBEVMSLLQPLNITQVLSHGEGTDINFWIQGVGA



Db 361 MESDAGFLPTGLOFTGSEKARAIMBEVMSLQPLNTQVLSHGEGDINFWIOAGVPGA 420  
 QY 421 SLDDLYKYFFFHSHGDTMTVMDPKOMNVAAAVAVSYVADMEEMLPRS 472  
 Db 421 SLDDLYKYFFFHSHGDTMTVMDPKOMNVAAAVAVSYVADMEEMLPRS 472

## RESULT 8

ABR48106  
 ID ABR48106 standard; Protein; 472 AA.

AC ABR48106;  
 XX 12-JUN-2003 (first entry)  
 DT XX

DE Human secreted protein, SEQ ID 997.

KM Cardiac; antiarrhythmic; antiarteriosclerotic; vasotropic; cyostatic;  
 KM vulnerable; antiinflammatory; nocrotic; neuroprotective;  
 KM antiparkinsonian; gene therapy; human; cardiovascular disorder.

OS Homo sapiens.

PN WO200295010-A2.

PD 28-NOV-2002.

PF 19-MAR-2002; 2002WO-US09785.

PR 21-MAR-2001; 2001US-277340P.

PR 19-JUL-2001; 2001US-306171P.

PR 13-NOV-2001; 2001US-331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;  
 PI WPI; 2003-129429/12.

PT Novel human secreted proteins, useful for detecting, preventing,  
 PT diagnosing, prognosticating, treating and/or ameliorating  
 PT cardiovascular disorders such as arrhythmia -

PS Claim 13; SEQ ID 997; 1881bp; English.

CC The present invention relates to novel human secreted proteins  
 CC (ABR47633-ABR48145) and their coding sequences (ACC50344-ACC50856). The  
 CC proteins and their coding sequences are useful for the preparation of a  
 CC diagnostic or pharmaceutical composition for diagnosing or treating a  
 CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,  
 CC coronary arteriosclerosis and myocardial ischemia), neural disorders,  
 CC immune system disorders, muscular disorders, reproductive disorders,  
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,  
 CC proliferative disorders and/or cancerous diseases and conditions, for  
 CC wound healing and epithelial cell proliferation, to treat inflammation or  
 CC infection, for treating thrombosis and arteriosclerosis, for treating or  
 CC preventing neural damage which occurs in neuronal disorders or  
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
 CC disease, to enhance bone and periodontal regeneration and aid in tissue  
 CC transplants or bone grafts, to prevent skin aging or hair loss, to  
 CC stimulate growth and differentiation of haematopoietic cells and bone  
 CC marrow cells when used in combination with other cytokines, to maintain  
 CC organs before transplantation or for supporting cell culture of primary  
 CC tissues, to increase or decrease differentiation or proliferation of  
 CC embryonic stem cells, or to modulate mammalian characteristics or  
 CC metabolism.

CC Note: The sequence data for this patent was published in electronic  
 CC format and is available from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 472 AA;  
 XX

Query Match 100.0%; Score 2417; DB 24; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 6,8e-212;  
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFPFEGCVHLISLCSGKAI CKNGISKRTTEETKEETASGCDVAKAIINLAVGKAQ 60  
 Db 1 MKFLIFAFPGVHLISLCSGKAI CKNGISKRTTEETKEETASGCDVAKAIINLAVGKAQ 60  
 QY 61 NRSYERLALVDVVPRLSSKNLEKAIQIMYONLQODGJEKYLEPVRI PHMERGESA 120  
 Db 61 NRSYERLALVDVVPRLSSKNLEKAIQIMYONLQODGJEKYLEPVRI PHMERGESA 120  
 QY 121 VMLPPIHKIAILGLSSISCTPEEGITAEVLVVTSPDELORASBARKIVVNPYINY 180  
 Db 121 VMLPPIHKIAILGLSSISCTPEEGITAEVLVVTSPDELORASBARKIVVNPYINY 180  
 QY 181 SRTVQYRTQGAVERAAKXGALASLIRSVASFSIYSPHTGIOEYDGVKPIPTACTIVEDAE 240  
 Db 181 SRTVQYRTQGAVERAAKXGALASLIRSVASFSIYSPHTGIOEYDGVKPIPTACTIVEDAE 240  
 QY 241 MMSRMASHGKIYIOLKMGAKTYPDTDSFNTVAETGSKYPEQVVLVSGHLSMDVGQGA 300  
 Db 241 MMSRMASHGKIYIOLKMGAKTYPDTDSFNTVAETGSKYPEQVVLVSGHLSMDVGQGA 300  
 QY 301 MDDGGAFISWEALSLIKDLRPRKTLRLVLTAEBOGGVGAFOYYQLHKVNI SNYSILV 360  
 Db 301 MDDGGAFISWEALSLIKDLRPRKTLRLVLTAEBOGGVGAFOYYQLHKVNI SNYSILV 360  
 QY 361 MESDAGFLPTGLOFTGSEKARAIMBEVMSLQPLNTQVLSHGEGDINFWIOAGVPGA 420  
 Db 361 MESDAGFLPTGLOFTGSEKARAIMBEVMSLQPLNTQVLSHGEGDINFWIOAGVPGA 420  
 QY 421 SLDDLYKYFFFHSHGDTMTVMDPKOMNVAAAVAVSYVADMEEMLPRS 472  
 Db 421 SLDDLYKYFFFHSHGDTMTVMDPKOMNVAAAVAVSYVADMEEMLPRS 472

## RESULT 9

ABR00152  
 ID ABR00152 standard; Protein; 472 AA.

AC ABR00152;  
 XX 03-APR-2003 (first entry)  
 DT XX

DE Human gene 142 encoded secreted protein HRA035, SEQ ID NO:441.

XX Human; secreted protein; digestive disorder; gastrointestinal disorder;  
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;  
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;  
 KW immune disorder; inflammation; infection; wound healing; drug screening;  
 KW chromosome identification; chromosome mapping; cytostatic; gene therapy;  
 KW antiinflammatory; immunosuppressive; vulnerable; chromosome 8q22.2.

OS Homo sapiens.

PN WO200276488-A1.

PD 03-OCT-2002.

PF 19-MAR-2002; 2002WO-US08276.

PR 21-MAR-2001; 2001US-277340P.

PR 19-JUL-2001; 2001US-306171P.

PR 13-NOV-2001; 2001US-331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;  
 PI WPI; 2003-029900/02.

XX N-PSDB; AB271331.  
 XX

PT New human secreted proteins and nucleic acids, useful for detecting,  
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating  
 PT e.g. gastrointestinal diseases and disorders, or cancers -  
 XX  
 PS Claim 13; Page 1029-1030; 1216pp; English.  
 CC AB271190-AB271478 represent cDNAs corresponding to 178 human secreted  
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.  
 CC AB271479-AB271540 represent human secreted protein genomic fragments. The  
 CC invention also encompasses antibodies specific for the secreted proteins,  
 CC the use of the secreted proteins in drug screening, and recombinant  
 CC vectors and host cells comprising a nucleic acid of the invention. The  
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
 CC fragments are useful for diagnosing, treating, ameliorating or preventing  
 CC digestive disorders. Such conditions include disorders of the mouth,  
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary  
 CC tract and pancreas, and include cancers of these organs and tissues. The  
 CC secreted proteins and their nucleic acids may also be used in the  
 CC treatment of immune disorders, inflammation, infection,  
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids  
 CC of the invention may be used for chromosome identification, chromosome  
 CC mapping, in gene therapy, for identifying individuals from minute  
 CC biological samples, as hybridisation probes, and as molecular weight  
 CC markers. The present sequence represents a human secreted protein of the  
 CC invention.  
 XX  
 SQ Sequence 472 AA;  
 Query Match 100.0%; Score 2417; DB 24; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-212;  
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAFGGVHLLSLCSGKAICNGISKRTPEEIKBEIASCGDVAKAIINLAAYGKAQ 60  
 DB 1 MKFLIFAFGGVHLLSLCSGKAICNGISKRTPEEIKBEIASCGDVAKAIINLAAYGKAQ 60  
 QY 61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHLEPRIPMERGESSA 120  
 DB 61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHLEPRIPMERGESSA 120  
 QY 121 VMLEPRIHKAIALIGLSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVVNOPIYNY 180  
 DB 121 VMLEPRIHKAIALIGLSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVVNOPIYNY 180  
 QY 121 VMLEPRIHKAIALIGLSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVVNOPIYNY 180  
 DB 121 VMLEPRIHKAIALIGLSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVVNOPIYNY 180  
 QY 181 SRTVOYRTOGAVEAAKVGALASLIRSVASFISYSPHTGIOEYQDGVPKIPACTIVEDAE 240  
 DB 181 SRTVOYRTOGAVEAAKVGALASLIRSVASFISYSPHTGIOEYQDGVPKIPACTIVEDAE 240  
 QY 181 SRTVOYRTOGAVEAAKVGALASLIRSVASFISYSPHTGIOEYQDGVPKIPACTIVEDAE 240  
 DB 181 SRTVOYRTOGAVEAAKVGALASLIRSVASFISYSPHTGIOEYQDGVPKIPACTIVEDAE 240  
 QY 241 MMSRHAHGKITIVIQKMGAKTYPTDSEFNTVAETISGKVPQVVLVSGHLDMSWVGGA 300  
 DB 241 MMSRHAHGKITIVIQKMGAKTYPTDSEFNTVAETISGKVPQVVLVSGHLDMSWVGGA 300  
 QY 241 MMSRHAHGKITIVIQKMGAKTYPTDSEFNTVAETISGKVPQVVLVSGHLDMSWVGGA 300  
 DB 241 MMSRHAHGKITIVIQKMGAKTYPTDSEFNTVAETISGKVPQVVLVSGHLDMSWVGGA 300  
 QY 301 MDDGGAFISWEALSLINDLGLRPRKTRLVMTAEQGGVGAFOYQOLHKVNISNYSVLV 360  
 DB 301 MDDGGAFISWEALSLINDLGLRPRKTRLVMTAEQGGVGAFOYQOLHKVNISNYSVLV 360  
 QY 301 MDDGGAFISWEALSLINDLGLRPRKTRLVMTAEQGGVGAFOYQOLHKVNISNYSVLV 360  
 DB 301 MDDGGAFISWEALSLINDLGLRPRKTRLVMTAEQGGVGAFOYQOLHKVNISNYSVLV 360  
 QY 361 MESDAGTLPYGLQFTGSEKARAIIEEYVSLIQPLINTQVLSHGEGTDINFWIOAGVGA 420  
 DB 361 MESDAGTLPYGLQFTGSEKARAIIEEYVSLIQPLINTQVLSHGEGTDINFWIOAGVGA 420  
 QY 421 SLDDLYVYFFPHSHGDTMTVMDBKQNVAAVAVVYVVAADMEEMLPSS 472  
 DB 421 SLDDLYVYFFPHSHGDTMTVMDBKQNVAAVAVVYVVAADMEEMLPSS 472

XX  
 DE Human gene 142 encoded secreted protein HRA035, SEQ ID NO:563.  
 XX  
 KW Human; secreted protein; digestive disorder; gastrointestinal disorder;  
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;  
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;  
 KW immune disorder; inflammation; infection; wound healing; drug screening;  
 KW chromosome identification; chromosome mapping; cytostatic; gene therapy;  
 KW antiinflammatory; immunosuppressive; vulnerary; chromosome 8q22.2.  
 XX  
 OS Homo sapiens.  
 XX  
 WO200276488-A1.  
 XX  
 03-OCT-2002.  
 XX  
 19-MAR-2002; 2002WO-US08276.  
 XX  
 21-MAR-2001; 2001US-277340P.  
 XX  
 19-JUN-2001; 2001US-306171P.  
 XX  
 13-NOV-2001; 2001US-331287P.  
 XX  
 (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM;  
 XX  
 WIPI, 2003-029900/02.  
 XX  
 N-PSDB; AB271453.  
 XX  
 PT New human secreted proteins and nucleic acids, useful for detecting,  
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating  
 PT e.g. gastrointestinal diseases and disorders, or cancers -  
 XX  
 PS Claim 13; Page 1108-1109; 1216pp; English.  
 CC AB271190-AB271478 represent cDNAs corresponding to 178 human secreted  
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.  
 CC AB271479-AB271540 represent human secreted protein genomic fragments. The  
 CC invention also encompasses antibodies specific for the secreted proteins,  
 CC the use of the secreted proteins in drug screening, and recombinant  
 CC vectors and host cells comprising a nucleic acid of the invention. The  
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
 CC fragments are useful for diagnosing, treating, ameliorating or preventing  
 CC digestive disorders. Such conditions include disorders of the mouth,  
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary  
 CC tract and pancreas, and include cancers of these organs and tissues. The  
 CC secreted proteins and their nucleic acids may also be used in the  
 CC treatment of immune disorders, inflammation, infection,  
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids  
 CC of the invention may be used for chromosome identification, chromosome  
 CC mapping, in gene therapy, for identifying individuals from minute  
 CC biological samples, as hybridisation probes, and as molecular weight  
 CC markers. The present sequence represents a human secreted protein of the  
 CC invention.  
 XX  
 SQ Sequence 472 AA;  
 Query Match 100.0%; Score 2417; DB 24; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-212;  
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAFGGVHLLSLCSGKAICNGISKRTPEEIKBEIASCGDVAKAIINLAAYGKAQ 60  
 DB 1 MKFLIFAFGGVHLLSLCSGKAICNGISKRTPEEIKBEIASCGDVAKAIINLAAYGKAQ 60  
 QY 61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHLEPRIPMERGESSA 120  
 DB 61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHLEPRIPMERGESSA 120  
 QY 121 VMLEPRIHKAIALIGLSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVVNOPIYNY 180  
 DB 121 VMLEPRIHKAIALIGLSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVVNOPIYNY 180

QY 181 SRTVYRTQGAWEAKVAGALASLRSVASFSIYSPHTGIOEYODGVPIPTACTIVEDAE 240  
 DB 181 SRTVYRTQGAWEAKVAGALASLRSVASFSIYSPHTGIOEYODGVPIPTACTIVEDAE 240  
 QY 241 MMSRMASHGKIYIVQLKMGAKATYPTDTSFNTVAEITGSKYEQVVLVSGHLSMDVGGGA 300  
 DB 241 MMSRMASHGKIYIVQLKMGAKATYPTDTSFNTVAEITGSKYEQVVLVSGHLSMDVGGGA 300  
 QY 301 MDDGGAFISWEALSLIKDGLRPRKTLRLVLTAEBOGGVGAPOYQLHKVNISNYSLV 360  
 DB 301 MDDGGAFISWEALSLIKDGLRPRKTLRLVLTAEBOGGVGAPOYQLHKVNISNYSLV 360  
 QY 361 MESDAGTFLPTGLOFTGSEKARAIEMEYMSLLQPLNTIQLVSHGEGTDINFWIOAGVPGA 420  
 DB 361 MESDAGTFLPTGLOFTGSEKARAIEMEYMSLLQPLNTIQLVSHGEGTDINFWIOAGVPGA 420  
 QY 421 SLDDLVYKYPFHSHSGDTMTVMDPKQNNVAAVAWVSYVADMEEMLPSS 472  
 DB 421 SLDDLVYKYPFHSHSGDTMTVMDPKQNNVAAVAWVSYVADMEEMLPSS 472

## RESULT 11

AA87081  
 ID AAY87081 standard; Protein; 473 AA.

XX AAY87081;

DT 09-MAY-2000 (first entry)

XX Human secreted protein sequence SEQ ID NO:120.

XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;  
 KM antiinflammatory; nootropic; neuroprotective; antiallergic; cancer;  
 KM tumour; neurodegenerative disorder; developmental abnormality; allergy;  
 KM foetal deficiency; blood disorder; immune system disorder; arthritis;  
 KM autoimmune disease; hepatic disease; renal disease; inflammation;  
 KM Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;  
 KM infection; AIDS; spinal cord injury; transplant rejection; diabetes;  
 KM asthma; sepsis; acne; psoriasis; cardiovascular disorder;  
 KM reproductive disorder; gastrointestinal disorder; respiratory disorder;  
 KM metabolic disorder; food additive; preservative.

OS Homo sapiens.

XX WO200004140-A1.

XX 27-JAN-2000.

XX 14-JUL-1999; 99WO-US15849.

XX 15-JUL-1998; 98US-0092921.

XX 15-JUL-1998; 98US-0092922.

XX 15-JUL-1998; 98US-0092956.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y,  
 PI Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE,  
 PI Mucenski M, Endress GA, Soppet DR;

XX MPI: 2000-161128/14.

XX N-PSDB; AA298034.

XX New isolated human genes, useful for diagnosis and treatment of, e.g.  
 PT cancers, neurological or blood disorders

XX Claim 11; Page 383-385; 494pp; English.

XX The polynucleotide sequences given in AA298017 to AA298108 encode the  
 CC human secreted proteins given in AAY87064 to AAY87223. Human secreted  
 CC protein can have activities based on the tissues and cells the genes are  
 CC expressed in. Examples of activities include: cytostatic;

CC immunosuppressive; antiinflammatory; nootropic; neuroprotective; and  
 CC antiallergic. The polynucleotides and their corresponding secreted  
 CC polypeptides are useful for preventing, treating or ameliorating medical  
 CC conditions, e.g. by protein or gene therapy. Also pathological conditions  
 CC can be diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new  
 CC polynucleotides. Human secreted protein s and their polynucleotides can  
 CC be used for developing products for the diagnosis or treatment of cancer,  
 CC tumours, neurodegenerative disorders, developmental abnormalities and  
 CC foetal deficiencies, blood disorders, diseases of the immune system,  
 CC autoimmune diseases, hepatic and renal disease, inflammation,  
 CC allergies, Alzheimer's disease, behavioural disorders, schizophrenia,  
 CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,  
 CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,  
 CC cardiovascular disorders, reproductive disorders, gastrointestinal  
 CC disorders, respiratory disorders and metabolic disorders. The  
 CC proteins or polynucleotides can also be used as food additives or  
 CC preservatives. The proteins are also useful for identifying their  
 CC binding partners. AA298008 to AA298016 and AAY87063 are sequence used in  
 CC the exemplification of the present invention.

SO Sequence 473 AA;

Query Match 100.0%; Score 2417; DB 21; Length 473;  
 Best Local Similarity 100.0%; Pred. No. 6,8e-212;  
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAPFGVHLLSLCSGKATCRNGISKTEFEIEBEIASCGDVAKAIINLAVYGAQ 60  
 DB 1 MKFLIFAPFGVHLLSLCSGKATCRNGISKTEFEIEBEIASCGDVAKAIINLAVYGAQ 60  
 QY 61 NRSYERLALVDYTPGRLSGSKNLEKAIQIWMYQLODGLKVLFPVRIPIHMERGESA 120  
 DB 61 NRSYERLALVDYTPGRLSGSKNLEKAIQIWMYQLODGLKVLFPVRIPIHMERGESA 120  
 QY 121 VMLBPRILHKAIIIGLGSSTGTPBGTAEVLVNTSPDELQRRASEARGKIVVYQPYINY 180  
 DB 121 VMLBPRILHKAIIIGLGSSTGTPBGTAEVLVNTSPDELQRRASEARGKIVVYQPYINY 180  
 QY 181 SRTVYRTQGAWEAKVAGALASLRSVASFSIYSPHTGIOEYODGVPIPTACTIVEDAE 240  
 DB 181 SRTVYRTQGAWEAKVAGALASLRSVASFSIYSPHTGIOEYODGVPIPTACTIVEDAE 240  
 QY 241 MMSRMASHGKIYIVQLKMGAKATYPTDTSFNTVAEITGSKYEQVVLVSGHLSMDVGGGA 300  
 DB 241 MMSRMASHGKIYIVQLKMGAKATYPTDTSFNTVAEITGSKYEQVVLVSGHLSMDVGGGA 300  
 QY 301 MDDGGAFISWEALSLIKDGLRPRKTLRLVLTAEBOGGVGAPOYQLHKVNISNYSLV 360  
 DB 301 MDDGGAFISWEALSLIKDGLRPRKTLRLVLTAEBOGGVGAPOYQLHKVNISNYSLV 360  
 QY 361 MESDAGTFLPTGLOFTGSEKARAIEMEYMSLLQPLNTIQLVSHGEGTDINFWIOAGVPGA 420  
 DB 361 MESDAGTFLPTGLOFTGSEKARAIEMEYMSLLQPLNTIQLVSHGEGTDINFWIOAGVPGA 420  
 QY 421 SLDDLVYKYPFHSHSGDTMTVMDPKQNNVAAVAWVSYVADMEEMLPSS 472  
 DB 421 SLDDLVYKYPFHSHSGDTMTVMDPKQNNVAAVAWVSYVADMEEMLPSS 472

## RESULT 12

AA87081  
 ID AAY87081 standard; Protein; 472 AA.

XX AAY87081;

DT 21-MAY-1998 (first entry)

XX Human secreted protein AM282 full-length sequence.

XX Secreted protein; AM282; cytokine; human.

OS Homo sapiens.

```

XX Key Location/Qualifiers
FH Peptide 1..24
FT /label= Sig_peptide
FT Protein 25..472
FT /label= Mat_protein
XX
XX WO9739030-A2.
XX
XX 23-OCT-1997.
XX
XX 16-APR-1997; 97WO-US06475.
XX
XX 13-JAN-1997; 97US-0783520.
XX
XX 18-APR-1996; 96US-0634325.
XX
XX (GENY ) GENETICS INST INC.
XX
XX Jacobs K, LaVallie ER, McCoy JM, Merberg D, Racie LA;
XX Spaulding V;
XX
XX MPI: 1997-526400/48.
XX
XX N-PSDB; AAV02296.
XX
XX New isolated secretory proteins AM340, AM282 and AK583 - possibly
XX have cytokine, cell proliferation/differentiation regulating,
XX immunomodulating activities, etc.
XX
XX Claim 16; Page 45-47; 59pp; English.
XX
XX This human secreted protein, designated AM282, is encoded by a
XX full-length cDNA clone (see AAV02296), deposited in ATCC 98026, that
XX was identified from a database search using an isolated partial
XX AM282 clone (see AAT97398). AM282 protein can be used in a claimed
XX method for preventing, treating or ameliorating a medical
XX condition. It may exhibit cytokine, cell proliferation (either
XX inducing or inhibiting) or cell differentiation (either inducing or
XX inhibiting) activity or may induce production of other cytokines in
XX certain cell populations. It may also exhibit e.g. immune
XX stimulating or suppressing activity, haematopoiesis regulating
XX activity, tissue growth activity, activin/inhibin activity,
XX chemotactic or chemokinetic activity, haemostatic or thrombolytic
XX activity, receptor/ligand activity, anti-inflammatory activity,
XX tumour inhibition activity, or other activities. No evidence of
XX any of these activities is given in the specification.
XX
XX
XX Sequence 472 AA:
SQ
Query Match 99.7%; Score 2410; DB 18; Length 472;
Best Local Similarity 99.8%; Pred. No. 3e-211;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKFLIPAFGGVHLLSLCSGKAIICKNGISKRTPEEIKKEIASCGDVAKAIINLVYGGKQ 60
DB 1 MKFLIPAFGGVHLLSLCSGKAIICKNGISKRTPEEIKKEIASCGDVAKAIINLVYGGKQ 60
QY 61 NRSYERLALVDVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESBA 120
DB 61 NRSYERLALVDVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESBA 120
QY 121 VMLERIRIKHIALIGSSIGTPPEGITAEVLVVTSTDELQRRASERKGIIVYNGPYINX 180
DB 121 VMLERIRIKHIALIGSSIGTPPEGITAEVLVVTSTDELQRRASERKGIIVYNGPYINX 180
QY 181 SFTVOYRTQGAVEAAKVGALASLIRSVASFSTIYSPHTGIOEVQDGVKIPITACTIVEDAE 240
DB 181 SFTVOYRTQGAVEAAKVGALASLIRSVASFSTIYSPHTGIOEVQDGVKIPITACTIVEDAE 240
QY 241 MMSRWASHSIKIVIOUKMGAKTYPPTDSFNTVAETGSKYPRQVNVLSHILDSMWVGCA 300
DB 241 MMSRWASHSIKIVIOUKMGAKTYPPTDSFNTVAETGSKYPRQVNVLSHILDSMWVGCA 300
QY 301 MDDGGAFISWEALSLIKDLGRPKRTLRVLMTAEQGGVGAFQYVQLHKVNISNYSIV 360

```

```

DB 301 MDDGGAFISWEALSLIKDLGRPKRTLRVLMTAEQGGVGAFQYVQLHKVNISNYSIV 360
QY 361 MESDAGTFLPTGLOFTGSEKRAIMEEYMSLLOPLINTOVLSHEGTDINFIQAGVGA 420
DB 361 MESDAGTFLPTGLOFTGSEKRAIMEEYMSLLOPLINTOVLSHEGTDINFIQAGVGA 420
QY 421 SLDDLYKYFFHHSHGDTMTVMDPKQMNVAAMVAVSYVAADMEMLPRS 472
DB 421 SLDDLYKYFFHHSHGDTMTVMDPKQMNVAAMVAVSYVAADMEMLPRS 472
RESULT 13
AAB10229
ID AAB10229 standard; Protein; 472 AA.
XX
XX AAB10229;
XX
XX 16-NOV-2000 (first entry)
XX
XX Human fetal kidney protein fragment AM282_1i.
XX
XX Secreted protein; cytosstatic; immunostimulatory; antimicrobial;
XX antiviral; immunosuppressive; antiinflammatory; vulnerary; cytokine;
XX cell proliferation; differentiation; regulator; treatment; tumor;
XX autoimmune disease; inflammatory disorder; wound; microbial infection;
XX viral disease; graft versus host reaction suppression.
XX
XX Homo sapiens.
XX
XX WO200037630-A1.
XX
XX 29-JUN-2000.
XX
XX 22-DEC-1999; 99WO-US31005.
XX
XX 23-DEC-1998; 98US-0220876.
XX
XX (GENY ) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
XX Merberg D, Treacy M, Bowman MR;
XX
XX MPI: 2000-442661/38.
XX
XX N-PSDB; AAA40493.
XX
XX Secreted human proteing AS296-1i and AS34-1i, useful for treating
XX tumors, autoimmune diseases, inflammatory disorders, wounds, microbial
XX infections and viral diseases -
XX
XX Disclosure: Page 199-200; 293pp; English.
XX
XX This invention describes novel secreted human proteins (I) which have
XX cytosstatic, immunostimulatory, antimicrobial, antiviral,
XX immunosuppressive, antiinflammatory and vulnerary activity and which act
XX as cytokine, cell proliferation or differentiation regulators. (I)
XX is useful for treating tumors, autoimmune diseases, inflammatory
XX disorders, wounds, microbial infections and viral diseases. (I) is also
XX useful for suppressing graft versus host reaction. AAB10226-B10288
XX represent the secreted proteins encoded by AAA40490-A40580 which are
XX described in the method of the invention.
XX
XX Sequence 472 AA:
SQ
Query Match 99.7%; Score 2410; DB 21; Length 472;
Best Local Similarity 99.8%; Pred. No. 3e-211;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKFLIPAFGGVHLLSLCSGKAIICKNGISKRTPEEIKKEIASCGDVAKAIINLVYGGKQ 60
DB 1 MKFLIPAFGGVHLLSLCSGKAIICKNGISKRTPEEIKKEIASCGDVAKAIINLVYGGKQ 60
QY 61 NRSYERLALVDVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESBA 120

```

```

Db      61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESA 120
Qy      121 VMLPRRIHKIALLIGSSIGTPPEGITAEVLVWTSFDELQRRASARKIIVVNOPTINY 180
Db      121 VMLPRRIHKIALLIGSSIGTPPEGITAEVLVWTSFDELQRRASARKIIVVNOPTINY 180
Qy      181 SRTVQYRTQGAVEAKVAGALASLIRSVAFSISYPHTGIEYODVPKIPACTIVEDAE 240
Db      181 SRTVQYRTQGAVEAKVAGALASLIRSVAFSISYPHTGIEYODVPKIPACTIVEDAE 240
Qy      241 MMSRMASHGIKIIVIQLMKGAATYPTDTSFNTVAEITTSKYPEQVVLVSGHLSMDVQGA 300
Db      241 MMSRMASHGIKIIVIQLMKGAATYPTDTSFNTVAEITTSKYPEQVVLVSGHLSMDVQGA 300
Qy      301 MDDGGGAFISWEALSLIKDGLRPRKRTLRLVMTAEBOGGVGAFOYQLHKVNISNYSLV 360
Db      301 MDDGGGAFISWEALSLIKDGLRPRKRTLRLVMTAEBOGGVGAFOYQLHKVNISNYSLV 360
Qy      361 MESDAGTFLPTGLOFTGSEKARAIMBEVMSLLOPLNTIQLVSHGEGTDINFWIQAGVPGA 420
Db      361 MESDAGTFLPTGLOFTGSEKARAIMBEVMSLLOPLNTIQLVSHGEGTDINFWIQAGVPGA 420
Qy      421 SLDDLYKYFFPFHSHGDTMTVMDPKOMNVAAMVAVSYVADMEMLPRS 472
Db      421 SLDDLYKYFFPFHSHGDTMTVMDPKOMNVAAMVAVSYVADMEMLPRS 472

```

## RESULT 14

AAM93559 standard; Protein; 472 AA.

AAM93559;

06-NOV-2001 (first entry)

Human polypeptide, SEQ ID NO: 3329.

Human, full length cDNA, cDNA synthesis; oligo-capping.

Homo sapiens.

EP1130094-A2.

05-SEP-2001.

07-JUL-2000; 2000EP-0114089.

08-JUL-1999; 99JP-0194486.

11-JAN-2000; 2000JP-0118774.

02-MAY-2000; 2000JP-0183765.

(HELI-) HELIX RES INST.

Ota T, Nishikawa T, Isogai T, Hayashi K, Ichii S, Kawai Y;

Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

WPI; 2001-524255/58.

N-PSDB; AAK94491.

830 Primers useful for synthesizing full length cDNA clones and their

use in genetic manipulation -

Claim 8, SEQ ID NO 3329, 1380bp + sequence listing, English.

The invention relates to primers for synthesizing full length cDNA

clones. 830 cDNA molecules encoding a human protein have been

isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA

molecules have been determined. Primers for synthesizing the full length

cDNA are useful for clarifying the function of the protein encoded by

the cDNA. The full length clones were obtained by construction of full

length enriched cDNA libraries that were synthesised by the oligo-capping

method. The primers enable the production of the full length cDNA easily

CC without any special methods. The present sequence is a polypeptide  
CC encoded by a full length human cDNA of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 472 AA;

Query Match 99.4%; Score 2402; DB 22; Length 472;  
Best Local Similarity 99.4%; Pred. No. 1.66-210;  
Matches 469; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1 MKFLIPAFGCVHLLSLCSGKAICKNGISKRTPEEIKELIASCGDVAKAIIINLAVYKAQ 60
Db      1 MKFLIPAFGCVHLLSLCSGKAICKNGISKRTPEEIKELIASCGDVAKAIIINLAVYKAQ 60
Qy      61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESA 120
Db      61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESA 120
Qy      121 VMLPRRIHKIALLIGSSIGTPPEGITAEVLVWTSFDELQRRASARKIIVVNOPTINY 180
Db      121 VMLPRRIHKIALLIGSSIGTPPEGITAEVLVWTSFDELQRRASARKIIVVNOPTINY 180
Qy      181 SRTVQYRTQGAVEAKVAGALASLIRSVAFSISYPHTGIEYODVPKIPACTIVEDAE 240
Db      181 SRTVQYRTQGAVEAKVAGALASLIRSVAFSISYPHTGIEYODVPKIPACTIVEDAE 240
Qy      241 MMSRMASHGIKIIVIQLMKGAATYPTDTSFNTVAEITTSKYPEQVVLVSGHLSMDVQGA 300
Db      241 MMSRMASHGIKIIVIQLMKGAATYPTDTSFNTVAEITTSKYPEQVVLVSGHLSMDVQGA 300
Qy      301 MDDGGGAFISWEALSLIKDGLRPRKRTLRLVMTAEBOGGVGAFOYQLHKVNISNYSLV 360
Db      301 MDDGGGAFISWEALSLIKDGLRPRKRTLRLVMTAEBOGGVGAFOYQLHKVNISNYSLV 360
Qy      361 MESDAGTFLPTGLOFTGSEKARAIMBEVMSLLOPLNTIQLVSHGEGTDINFWIQAGVPGA 420
Db      361 MESDAGTFLPTGLOFTGSEKARAIMBEVMSLLOPLNTIQLVSHGEGTDINFWIQAGVPGA 420
Qy      421 SLDDLYKYFFPFHSHGDTMTVMDPKOMNVAAMVAVSYVADMEMLPRS 472
Db      421 SLDDLYKYFFPFHSHGDTMTVMDPKOMNVAAMVAVSYVADMEMLPRS 472

```

## RESULT 15

AAM25762 standard; Protein; 481 AA.

AAM25762;

16-OCT-2001 (first entry)

Human protein sequence SEQ ID NO:1277.

Human, cancer; ulcer; HIV infection; human immunodeficiency virus;

antimicrobial; antineoplastic; antiarthritic; immunosuppressive;

antibacterial; endocrine; cardiac; central nervous system; vitruide;

anti-HIV; fungicide; antimutagen; cardiovascular; antihaemic; anaemia;

antitumor; haemostatic; vulnery; antiulcer; osteoporosis; eczema;

dermatological; antiallergic; antidiabetic; cytostatic;

neuroprotective; nootropic; antiparkinsonian; infection;

immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;

antiaphyllactic; rheumatoid arthritis; septic shock; pancreatitis;

cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;

genetic disease; haematopoietic disorder; platelet disorder; asthma;

chromocytopenia; osteoporosis; severe combined immunodeficiency;

allergic rhinitis; diabetes; multiple sclerosis; depression;

Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;

neurological disorder.

Homo sapiens.

WO200153455-A2.

XX 26-JUL-2001.  
PD 22-DEC-2000; 2000MO-US35017.  
XX  
XX 23-DEC-1999; 99US-0471275.  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
XX  
PA (HYSE-) HYSEQ INC.  
PI Tang YT, Liu C, Drmanac RT;  
XX WPI; 2001-457603/49.  
DR N-PSDB; AAH99703.  
XX  
PT Isolated human polynucleotides encoding polypeptides, useful for the  
XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection -  
XX  
PS Claim 20; Page 265; 1217p; English.  
XX

CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
CC AAM25963. The proteins can have activities based on the tissues and  
CC cells they are expressed in, such as: antiinflammatory; antineumatic;  
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;  
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;  
CC cardiovascular; antianaemic; antidiabetic; haemostatic; vulnery;  
CC antitumor; osteopathic; dermatological; antiallergic; antiasthmatic;  
CC antidiabetic; cytosolic; neuroprotective; antidepressant; nootropic;  
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
CC encoding them can be used in gene therapy, antisense therapy and vaccine  
CC production. The proteins and polynucleotides are useful for screening for  
CC agonists or antagonists of a protein and for the treatment and diagnosis  
CC of disorders associated with the activity of a protein e.g. inflammation,  
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
CC anaemia, platelet disorders, thrombocytopaenia, wounds, ulcers,  
CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
CC neurological disorders.  
XX  
XX

SQ Sequence 481 AA;

Query Match 98.4%; Score 2378.5; DB 22; Length 481;

Best Local Similarity 98.9%; Pred. No. 2.3e-208;

Matches 468; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 MKFLFAFFGGVHLISLCSGKAIICNGISKRTFEIEKEIASCGPAKAIINLAAYGKAQ 60  
DB 5 MKFLFAFFGGVHLISLCSGKAIICNGISKRTFEIEKEIASCGPAKAIINLAAYGKAQ 68  
QY 61 NRSYERLALVDVTGPRISGSNLEKAIQIMYQNLQODGLEKHLPEVRIPIHMERGESA 120  
DB 65 NRSYERLALVDVTGPRISGSNLEKAIQIMYQNLQODGLEKHLPEVRIPIHMERGESA 128  
QY 121 VMLERIRIKIALIGSSISGTPPEGITAEVAVVTSFDELQRRASERAGKIVVYNQPIYNY 180  
DB 129 VMLERIRIKIALIGSSISGTPPEGITAEVAVVTSFDELQRRASERAGKIVVYNQPIYNY 188  
QY 181 SRTVOYRTQGAVEAKVGLASLIRSVASFISYSPHTGIOEYQDGVPKIPTACITVEDAE 240  
DB 189 SRTVOYRTQGAVEAKVGLASLIRSVASFISYSPHTGIOEYQDGVPKIPTACITVEDAE 248  
QY 241 MMSRNAHSGIKIVIQKMGAKTYPDTDSFNTVAIBITGSKYRBOVVLVSGHLSMDVQGA 300  
DB 249 MMSRNAHSGIKIVIQKMGAKTYPDTDSFNTVAIBITGSKYRBOVVLVSGHLSMDVQGA 308  
QY 301 MDDGGGAFISWALSLIKDLGIRPKRTLRVLMVTAEBQGVGAFQYQQLHKVNISNYSIV 360  
DB 309 MDDGGGAFISWALSLIKDLGIRPKRTLRVLMVTAEBQGVGAFQYQQLHKVNISNYSIV 368

QY 361 MESDAGTFLPTGLQFTGSEKARAIMERYMSLLOPLNITOVLSHSGEGTDINFIQAGVGA 420  
DB 369 MESDAGTFLPTGLQFTGSEKARAIMERYMSLLOPLNITOVLSHSGEGTDINFIQAGVGA 428  
QY 421 SLDDLYKYFFFHSHSGDPTMTMD-PRQMNVAALVAVSVYVADMEBMLPRS 472  
DB 429 SLDDLYKYFFFHSHSGDPTMTMHGIQTQMNVAALVAVSVYVADMEBMLPRS 481

Search completed: December 22, 2003, 11:49:03  
Job time : 82.2473 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:37:50 ; Search time 38.6646 Seconds  
(without alignments)  
1173.984 Million cell updates/sec

Title: US-09-745-763-36

Perfect score: 2417  
Sequence: 1 MKFLIPAFGCVHLISLSCG.....AAVAVSVYVADMEMLPRS 472

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_76:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	759.5	31.4	493	2 T46974	leucyl aminopeptid
2	640	26.5	467	2 G87564	aminopeptidase, pr
3	182	7.5	536	2 B83278	probable aminopept
4	176.5	7.3	571	2 H87600	hypothetical prote
5	173.5	7.2	776	1 JH0570	transferrin recept
6	169	7.0	455	2 S39663	aminopeptidase hom
7	165.5	6.8	703	2 T47631	peptidase-like pro
8	165	6.8	466	2 B84132	aminopeptidase B3
9	165	6.8	750	2 A56881	prostate-specific
10	163	6.7	1483	2 T19751	hypothetical prote
11	161.5	6.7	751	2 T30154	hypothetical prote
12	158.5	6.6	501	2 G82414	aminopeptidase VCA
13	156	6.5	433	2 P84215	aminopeptidase lim
14	150.5	6.2	529	2 G82759	hypothetical prote
15	149.5	6.2	609	2 JG7819	metalloproteinase
16	149	6.1	500	2 H70629	probable AMINOPEPT
17	147.5	6.1	441	2 A84351	hypothetical prote
18	147.5	6.1	504	2 S24314	bacterial leucyl a
19	146	6.0	430	2 A87634	peptidase, M20/M25
20	143.5	5.9	622	2 A34549	transferrin recept
21	140	5.8	488	2 H90220	hypothetical prote
22	138.5	5.7	284	2 S66427	aminopeptidase (EC
23	138.5	5.7	374	2 S69639	hypothetical prote
24	137.5	5.7	537	2 A54134	aminopeptidase y (
25	129.5	5.4	811	2 B57149	probable membrane
26	129	5.3	280	2 B5075	probable peptidase
27	126	5.2	760	1 UJHU	transferrin recept
28	125.5	5.2	413	2 S45467	aminopeptidase y h
29	122	5.0	757	2 A48592	transferrin recept

30	121.5	5.0	488	2 A87569	peptidase M20/M25/
31	117.5	4.9	1838	2 H86337	protein F5M15.26 (
32	114	4.7	474	2 E87650	peptidase, M20/M25
33	114	4.7	763	1 S29548	transferrin recept
34	113.5	4.7	794	2 T40289	hypothetical prote
35	111	4.6	411	2 A10567	allantoin amidohy
36	106.5	4.4	557	2 G72356	penicillin-binding
37	105.5	4.4	470	2 S57902	peptidase V - Lact
38	105.5	4.4	486	2 A82427	pyruvate kinase II
39	105	4.3	411	1 C64783	probable N-carbamyl
40	105	4.3	450	2 S43914	hypothetical prote
41	103.5	4.3	933	1 B48349	glycoprotein B pre
42	102	4.2	1332	2 F69732	PBSX prophage ORF
43	101.5	4.2	336	2 C83926	S-adenosylmethion
44	101.5	4.2	411	2 B71451	probable desucciny
45	101.5	4.2	470	2 AD1277	Xaa-His dipeptidase

#### ALIGNMENTS

##### RESULT 1

T46974 leucyl aminopeptidase (EC 3.4.11.-) ES-62 precursor [validated] - nematode (Acanthocheil, leucyl aminopeptidase (EC 3.4.11.-) ES-62 precursor [validated] - nematode (Acanthocheil, C/Species: Acanthocheilonema viteae C/Date: 17-Mar-2000 #sequence\_revision 17-Mar-2000 #text\_change 01-Feb-2002 C/Accession: T46974 R/Harnett, W.; Houston, K.M.; Tate, E.; Garate, T.; Apfel, H.; Adam, R.; Haslam, S.M.; Pa Mol. Biochem. Parasitol. 104, 11-23, 1999 A/Title: Molecular cloning and demonstration of an aminopeptidase activity in a filarial A/Reference number: Z24355, PMID:20055923; PMID:10589978 A/Accession: T46974 A/Status: preliminary; translated from GB/EMBL/DBJ A/Molecule type: mRNA A/Residues: 1-493 <HR> A/Cross-references: EMBL:AF071194; NID:G3386477; PIDN:AA28365.1; PID:G3386478 C/Function: A/Description: EC 3.4.11.- leucyl aminopeptidase [validated, PMID:20055923] A/Keywords: aminopeptidase

Query Match	31.4%	Score 759.5;	DB 2;	Length 493;
Best Local Similarity	36.6%	Pred. No. 2e-50;		
Matches 174;	Conservative 92;	Mismatches 166;	Indels 43;	Gaps 13;
QY	15	LSLCSGKALC-----KNGISKRTPEIKKEELASGCVAKAIINLAVGKAQNSYERL	67	
DB	14	LTVALGAALVDPDKTVAAPKNYIOETFGKVAE-----LIQYITKGEVGLAYQWL	62	
QY	68	ALAVDTVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLLEPV-RIPMERGESAVMLEPR	126	
DB	63	SKLVDFGHHMVGSDSLKSIATLESLKNDNDKHTTEVPNLPVHVRKNDVEMLEPR	122	
QY	127	IKHIAIIGLSSIGTPEEGITAEVLVVTSPDELQRRASEARSKIYVNYQPYNSRTVOY	186	
DB	123	NORLNVLAIG---GSEBASATGEVTVYDDDV--KDDVRGKIIVTAQFAGPYLTKY	177	
QY	187	RTQGAVERAAVGAALASLRVASFSIYSPHTGQIOYDQIPKIPACTITYEDAEKSRMA	246	
DB	178	R-BSVKLFQDLAGIVLKSITFSINSPTTGAEN--TTIPAACTITEEMERLY	233	
QY	247	SHGKIYIOLKMGAKTPTDTSFNTVAETIGSKYKPEOVVLVSGHLSMDVGGQAMDGGG	306	
DB	234	RSGKVIYIRMDKSHYEPINSSNLRIETIGSERPSEVLLSAHVDSMDVGGQALDDGAG	293	
QY	307	AFISWEALSLIKDGL---LRPKETRLVLTAEQGGVGAFOYQLHKVNI-SNYSLV	360	
DB	294	CAVWMSALHSLIKLAERNPKFKPKRTIRGIFMTSEEGYGAGKHYYITHKNDPEKEFYFV	353	
QY	361	MESDAGFLPLPTG----LQFTGSEKARIMEWMSLQPLNITQVL--SHBEGDINWIQ	414	
DB	354	SEIDTGTFFKSTNMLAHLSFSGDKSKMLRKEITRLLSRNGIAGLINSVQGG-DVTWAK	412	
QY	415	AGVVGASGLDD--LYKYFFHHSHGDTMTWMDPKQNNVAAVAVSVYVADME 467		

Db 413 DGIPSVNYIPDKAVDYFFYHHHTAGDYMTVLKQDLEYYTTSIFALIGHVIANMDD 4673

## RESULT 2

687564  
 aminopeptidase, probable [imported] - *Caulobacter crescentus*  
 C:Species: *Caulobacter crescentus*  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C:Accession: G87564  
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonitskiy,  
 n, J.; Armitage, M.; White, O.; Salzberg, S.L.; Shapiro, V.; Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: G87564  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-467 <STO>  
 A:Cross-references: GB:AE005673; NID:gl3424107; PIDN:AAK24515.1; GSPDB:GN00148  
 C:Genetics:  
 C:Gene: CC2544

Query Match	26.5%	Score 640;	DB 2;	Length 467;
Best Local Similarity	36.2%;	Pred. No. 2.8e-41;		
Matches 147;	Conservative 72;	Mismatches 177;	Indels 10;	Gaps 4

```

Oy 70 VVDVTPRPLSGSKNLEKAIQIMYONULQODGLEKHYLEPRIIHWMEGESSAAVMLEPRIHK 129
Db 49 LTTNIGPRIVSGPAAAKADWSVAKFKALGFTNNIKIDEPFAKPSWSGSEBSAEIUVAYAMK 108
Oy 130 IALIGLSSIGTPREGITAEVLVWTSFDELQIRASEA-RGKIUVWYQPYINYSRTVOY-- 186
Db 109 LGAVELGTVSTIRPAGIGIEVALFTKIPADMAAPDALKGKIIVITQPNVTRQNGAGYA 168
Oy 187 ---RTQGAVEAKVGALASLIRSVASFISYSPHTGISOEYODGVKPIPPACTIVBEDAEMK 242
Db 169 AGVSRIRIPVEBAKAGAVAMLIRISISTSDSTVPHGTGTSAGGVVTIPAAAGVPEAQUL 228
Oy 243 SRMAHGIKIVIQVCKMGAKTYPRDDBSFNVVAITGSKYPEQVNVLSGCHLDDSMVDGGMAD 302
Db 225 ERLAARKPMPRYKLKESSVNNNNVAMNISGGIKSSEKDEIVYIGGHLDSMDVGTGALD 288
Oy 303 DGGGAFISWEALSLIKDGLRPRKRTILVLTAEBOGGVGAFOYQYOHKYNISNYSVLWE 362
Db 289 DATGIAITTAAKKILGDLRPRKRTIRVVMKGSSEBGS--SSEYVLAANDALSTWVLAE 347
Oy 363 SDAGTFLPTGLQFTSGSEKARALMEWMSLLQPLNI--TQVLSHGEGTDINFWIQAGVGA 420
Db 348 SDTGADRTIYSLQVPGSLDHPVYVKAASVLAELPKIIVDRTPAAGAGADVSGIERAGVPYI 407
Oy 421 SLDDLKYKFFPHSHSGDTMTWMDPRKONVAAAANAVALVSYVVADE 466
Db 408 NMQDASRYFDYHHTTMDTLIDKYRPAELKQONVAAMWLSLVLAADSO 453

```

### RESULT 3

```

B813278
probable aminopeptidase PA2939 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B813278
S:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
adman, S.; Yank, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kss, A.; Ladbis, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10994043
A:Accession: B813278
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-536 <STO>

```

A:Cross-references: GB:AE004720; GB:AE004091; NID:g99499032; PIDB:AAG06327.1; GSPDB:GN0017  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA2939

Query Match 7.5%; Score 182; DB 2; Length 536;  
 Best Local Similarity 23.6%; Pred. No. 6e-06;  
 Matches 95; Conservative 61; Mismatches 141; Indels 106; Gaps 19;

```

QY 65 RRLLLVDTVPBRLTSGSSNNLEKALQIWMQNLQDQGLEVHLEPRIPRMWEGEEAAVMLE 124
Db 75 EDISLND--GNRRAATPGQASVDYKQTLQKAGY-KVSQVPPFTAY-----Y 121
QY 125 PRIKIALILGSSIGTPPE-----GITAELVAV-----TS 155
Db 122 PK-----GGSLSATVPQPVTYEMEKDFTYLSQTEAGDVTAKVPPVDDLGLAGNTSTS 174
QY 156 FDELQRRASRARGKI VYVNOPIYNSRTVOYRTQGAVEAKVGLASLIRSVAESISYSP 215
Db 175 GCEAEDEFANFPAGSIALIQRGTGNFE-----QKAENMAAAGAAGVITIFNOGN---TDD 224
QY 216 HTGL-----QBYQDVPKIPTACTIYEDAEEMSRMASHGKIYIQLKMGAKYTPDITDSF 269
Db 225 RKGJENVTGVSYSYGGLFVLI---FATJDNQVAWSQTPDLQJLHLVDV-----VRKRIETY 276
QY 270 NTVAIEITGSKYPEBOVLVSGHLSMDVQOGAMDQGGAFISME-ALSLIKDILGRPKRTL 328
Db 277 NVVAIE-TRGRPNNVVMVGAHLDSVFBSPGJNDNGSSGAOLQENAVLLAKL--PVUKV 332
QY 329 RLVLMTAEEOGVGAFOYYQLHKVNI SNYSLVMSDAQTFPLPTGLQPTGSEKARAMEBV 388
Db 333 RFAMWGAEBALVQSTHYVQ-----NLAPEE-----KKIKIAYLN-- 367
QY 389 MSLIOLPILITOVLSHGEGTDINFWIQAGVPGASLIDLIDYKTYF 431
Db 368 FDMIGSPFNFGNIYDQGSDD--FGIQ-GPPSSAIEIRLFEBEYF 407

```

## RESULT 4

H87600  
hypothetical protein CC2840 [imported] - Caulobacter crescentus  
C|Species: Caulobacter crescentus  
C|Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #ext\_change 10-May-2001  
C|Accession: H87600  
R|Niemann, W.C.; Feldbljum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Duxkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A|Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A|Reference number: A87249; MUID:21173698; PMID:11259647  
A|Accession: H87600  
A|Status: preliminary  
A|Molecule type: DNA  
A|Residues: 1-571 <STO>  
A|Cross-References: GB:AB005673; NID:gl3424450; PIDN:AAK24804.1; GSPDB:GN00148  
C|Genetics:  
A|Gene: CC2840  
C|Superfamily: Xylella fastidiosa hypothetical protein XF0820

Query Match	7.3%	Score 176.5;	DB 2;	Length 571;
Best Local Similarity	23.9%;	Pred. No. 1.8e-05;		
Matches 97;	Conservative 60;	Mismatches 132;	Indels 117;	Gaps 20

```

QY      129 KIALIGLOSSIGTPEGLTAELVWTSDELQRBASERKGIWY-----NQPV- 177
Db      180 KIIVW-----LGGAPASIQTERRAHLSPNTRRAEERKGIWVITLSTPGEKKRRPKA 234
QY      178 -----INVSRTQVYRTQ--GAVEAAKVGELASLIRSASFISYPHTGICQEVODGVKPIPT 231
Db      235 GVAGMKSRVYWRNAQDVGAIRAACTPELASLSQAGAA-KLPAGAG-----SLET 284
QY      232 ACTIVEDAEMNSRMAISHGIKIVIQLKMGAKT--YPTDTSFNTVAETGS--KYPEQVLIWS 288

```







A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-750 <ISR>  
 A:Cross-references: GB:999487; NID:g190663; PIDN:AAA60209.1; PID:g190664  
 A:Experimental source: prostatic carcinoma cell line LNCaP  
 A>Note: sequence extracted from NCBI backbone (NCBI:121724, NCBI:121725)  
 C:Superfamily: transfeitin receptor  
 C:Keywords: surface antigen; transmembrane protein

Query Match 6.8%; Score 165; DB 2; Length 750;  
 Best Local Similarity 19.2%; Pred. No. 0.0002;  
 Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

```

Qy 3 FLTFAFGVHLLSLCSGKAI CKNGISKRTFEI KEELIASCGDVAKAIINLAVYGAQNR 62
D 32 FLLGLFLEGMFKSNENATNTTPKHNKAFDELKAE-----NKKFLVNF----- 77
Qy 63 SYERLLLVDTVGRPLSGSK--NLEKAIQIMYQNLQODGLEKVLBEV-----RIPHW 113
D 78 -----TQIDPLAGTEQNQLAKQIQGQKKEFGDVELAHVDLLSYPKTHPNY 127
Qy 114 -----ERGEE--SAVMEPRIRHKAIILGLSSIGTPPGITAEVLVVTSPDELQRRASEA 166
D 128 ISIIINEGNEIFNLSLEP-----PPGYENVSDIVPPGSAFSPQCM-P 170
Qy 167 RGIIVVNOPPIYNSRT-----VOYRTQGAWEAKVG-----ALASLIRSV 207
D 171 EGDLY-----YVYAKTEDPFLERDMKINSGKIVIAKYGRKVKVKAQLAGAKGV 225
Qy 208 ASFS-----IYSPHTGIOEYOD-----ALASLIRSV 224
D 226 ILVSDPADYFAP--GVKSYPDGMNLPGGGVQGRNINLNGADPLTPGYPANEAAYRRGI 283
Qy 225 -----GVKRIPIACTVDEAEKMSMAHSIGITIVQLKMGATTPD----- 265
D 284 AEAVGLPSIPVHPHIGYDDAQLLE-----KVGSGAPDSSWRGSLKVPYVNGP 331
Qy 266 -----TDSFNTVAEITGSKYPRQVVLVSGHLSMDVMDVGQAMDG 304
D 332 GTFNQSTQKVKMHIHNEVTRIVYVIGTLRGAVEPRVYIILGHRDSWVF--GGIDPQ 389
Qy 305 GGAFISWE--ALSLIKDLGLRPKRTLLVMTAEEOGVGAFOYQLHKVNISNYSIVM 361
D 390 SGAIVVHIVRSFGLKKEGWRPRTILFASMDABEFLGSTWAE-----ENSRLLQ 443
Qy 362 ESDAGTFLPTGLQFTGSEKARAIMBEVW-SLLQPLNTIQTUVRSHGCDINFIWQAVPGA 420
D 444 ERGV-AYINADSIEGNYTLFVDCITPLMYSLVH--NLTKELKSPD-----EGPEGK 491
Qy 421 SLDDLKX 428
D 492 SLYESWTK 499

```

RESULT 10  
 T19751  
 hypothetical protein C35C5.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
 C:Accession: T19751  
 R:White, S.  
 submitted to the EMBL Data Library, August 1996  
 A:Reference number: Z19173  
 A:Accession: T19751  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1483 <MTL>  
 A:Cross-references: EMBL:Z78417; PIDN:CAB01688.1; GSPDB:GN00028; CESP:C35C5.2  
 A:Experimental source: clone C35C5  
 C:Genetic8:  
 A:Gene: CESP:C35C5.2  
 A:Map position: X  
 A:introns: 61/1; 203/3; 248/1; 293/3; 331/1; 437/1; 485/2; 523/2; 556/1; 595/3; 620/2; 6

Query Match 6.7%; Score 163; DB 2; Length 1483;  
 Best Local Similarity 19.5%; Pred. No. 0.0003;  
 Matches 108; Conservative 82; Mismatches 181; Indels 184; Gaps 24;

```

Qy 25 KNIGSKRTFEI EEE--IASCGDVAKAIINLA--VYGAQNRSYE-----RLALLVD 72
D 7 KHTFLQKVFQSKSRNKGKASGVAVLAVSTALTIILSNAIHOSYKSNKFLPKLSIATV 66
Qy 73 -----TVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLBEVPRP-----P 111
D 67 IKOLINVDNIRSNLHSLTKKPPHVAQTENNLRVEMIRDOITQGLEVHNEYNMLSYSP 126
Qy 112 HMERGESAVMLPRIRHKAIILGLSSIGTPPGITAEVLVVTSPD---ELQRRASEARG 168
D 127 NW-----TTPNT--IILKENDDLVSTTGRSVVKEQNDRPLAETIOWLAYSAG 175
Qy 169 KI---VVY-----NOPPIYNSRTV-----OYRTQGAWEAKVAGLASLI-- 204
D 176 TVEGDIYVNNANPSDIEYLESIGIDLKDKIFLARYSSNYRGNIAQMAVKGAACLVYS 235
Qy 205 --RSVASF-----SIY-----SPTGT-----OEYQ 223
D 236 DPRQVASLGTGPNETYGNTDKMPSNTVQMGSVYIGLDPPTPAFSGIDLFKETEGDDL 295
Qy 224 D--GVKPIPIACTVDEAEKMSMAHSIGITIVQLKMGATTPD-----SRMASHG 249
D 296 DEKPIPIPIACTVDEAEKMSMAHSIGITIVQLKMGATTPD-----SRMASHG 249
Qy 250 -----IKIVIQLMGAKTYPDTSFNTVAEITGSKYPRQVVLVSGHLSMDVQGA 300
D 356 ENEBSRIVVELLSRLRLNFRKIO-----NIMGVIKSGQSPDRFVLVSNHYDAWT--YGA 408
Qy 301 MDGCGAFIWEALSILK---DLGLRPKRTLLVMTAEEOGVGAFOYQLHKVNISN 356
D 409 VDPNSGTSTLLEVSRAIKQYQNTGWIIPARSILFAMHDAEYGLIGSTFAEEYRLQUMR 468
Qy 357 YSL-VNESDAGTFLPTGLQFTGSEKARAIMBEVW-SLLQPLNTIQTUVRSHGCDINFIWQ 415
D 469 RAVAVIMMDLIGAGNQTLLIGISNFTVANVLASAANVQF--NPTBM----- 512
Qy 416 GVPGASLDDLKXKF 430
D 513 -EGRKTLVDSWKY 526

```

RESULT 11  
 T30154  
 hypothetical protein R57.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
 C:Accession: T30154  
 R:Favell, T.; Rifkin, L.; Chiapelli, B.  
 submitted to the EMBL Data Library, February 1997  
 A:Description: The sequence of C. elegans cosmid R57.  
 A:Reference number: Z20745  
 A:Accession: T30154  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-751 <FAV>  
 A:Cross-references: EMBL:U88179; PIDN:AB52660.1; GSPDB:GN00028; CESP:R57.1  
 A:Experimental source: strain Bristol N2; clone R57  
 C:Genetic8:  
 A:Gene: CESP:R57.1  
 A:Map position: X  
 A:introns: 40/1; 145/1; 230/1; 275/3; 313/1; 344/2; 405/1; 453/2; 576/2; 671/2; 713/3

Query Match 6.7%; Score 161.5; DB 2; Length 751;  
 Best Local Similarity 18.8%; Pred. No. 0.00038;  
 Matches 118; Conservative 108; Mismatches 188; Indels 213; Gaps 33;  
 4 LIPAFPGVHLLSLCSGKAI CKNGI--SKRTFEI KEELIASCGDVAKAIINLAVYGA 58

Db 12 LIEVF-----CIAALGVHSEKRNKFNK--VSIDIDHKSADAGVIOONI--K 54  
 QY 59 AON-ASYERLALVDTVPRLSGSKULEKAIQIMYONLQODGLEKYLEPVRNI-----PHM 113  
 Db 55 TENIKKY-----LRITKDPHVAGTEANKVAVETIANAMSEAGLEDVHTLPYVLLSYDPF 110  
 QY 114 E-----RGEESAVMLEPRIRK-----IAIILGSSIGTPREGITAE 149  
 Db 111 ENPNISVIIKSSAGKEVFKSGVSPVILPDEQSGKYAGHQWLAIVAGNSA-----SAD 162  
 QY 150 VLVV-----TSFDELQORASEARKIVVYNQPIYNSRVQRTQGAVEAAVGAALASLI 204  
 Db 163 VYIINHGTANDPKNLKMGVDIKGKIAL-----MRYGH--GFRGDKIKHNAQGAIGAIL 215  
 QY 205 RS-----VASEFSIY-----SPHTGIOE----- 221  
 Db 216 FSDTODVADGVSESENVYPKKIMWPEGVQSGSLMHGDDALSPPYPSKELFKGRTIEE 275  
 QY 222 -YQDGV-PKIPACTIVEDA-EMMSRMAS-----HGKIKVI 254  
 Db 276 AKEDGVLPSPVLPVPSYTTGYEILKRLSGRAPSPDQGVGVGNLTYYKLGPFVNGEKLSI 335  
 QY 255 QLKMAKTYPTDSEFNTVAETGSKTPEQVVLVSGHLSWDVQGGAMDGGGAFISWEA- 313  
 Db 336 NVHSELRT--KRIRNVIGYIRGSEBPSYIMLGHHFDM--VGSIDPNSGTAVALAEVA 390  
 QY 314 ---LSLIKDLGLRPKRTLVLVMTAEQGVGAFOYOLHKVNI-----S 355  
 Db 391 RAMQOTINETSMPKPARTIYFNAMDAEFPGLISTEVE-EFVNIIQKRAVVTYINMDCIOG 449  
 QY 356 NYSLVMS-----DAGTELPGLQFTGSEKARAIEMEWSLL--QPLNITQVLSHG 404  
 Db 450 NLSLHVDTVPILIEHVAIEASKOVENPSKRSRGRKTLVDYTMKVPFPDKACVPRIRVPG 509  
 QY 405 ESTD-----INFYIQGVCPASLDDLYKFFPHSHGDT-----MTVMPK 446  
 Db 510 GGSIDAPPLNF--AGVP-----VINFTFKYTTWTDYPLHYTMETPFPSNHLPLTD 559  
 QY 447 QMNVAAAV--WAVSYVVADEMEMLP 470  
 Db 560 NLSYKALIGQYMAELAKTPAD-DVILP 585

RESULT 12  
 C82414  
 antipeptidase VCA0813 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
 C.Species: Vibrio cholerae  
 C.Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C.Accession: C82414  
 R.Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.  
 1. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A.Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A.Reference number: A82035; MUID:20406833; PMID:10952301  
 A.Accession: C82414  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-501 <HEI>  
 A.Cross-references: GB:AE004409; GB:AE003853; NID:g9658233; PIDN:AAF96711.1; GSPDB:GN001  
 C.Genetics:  
 A:Gene: VCA0813  
 A:Map position: 2

Query Match 6.6%; Score 158.5; DB 2; Length 501;  
 Best Local Similarity 21.4%; Pred. No. 0.00035;  
 Matches 75; Conservative 63; Mismatches 139; Indels 73; Gaps 12;

QY 168 GRIVYVYNQPIYNSRVQRTQGA--VEAAKVGALASLI--RSVASFSIYSPHTGIOEXQ 223  
 Db 56 GEVALDELAELASHTWHEOHNRGCGYVWHTSAGAMAAALMMPESIANFTIIPAD-----SQQ 110

QY 224 DGV-----PKIPACTI-----TVDAEMSMRASHGIKIVIOI--KXGA 260  
 Db 111 DLVNAVLPOVADQITVTIRALSSFNRRFYTTASGAQSDMLANEMSLISLPGSRLEQ 170  
 QY 261 KTYPTDSEFNTVAETGSKTPEQVVLVSGHLS-----WDVQGGAMDGGGAFISWE 312  
 Db 171 IKHSGVYKQSVVLTIOSEKDEWYIVGCHLDSITLGSHTNEOSTAPGADDASGIALSE 230  
 QY 313 ALSLIKDLGLRPKRTLVLVMTAEQGVGAFOYOLHKVNIYSVLSMESDAGTFLPTG 372  
 Db 231 IIRVLRRNRRFPKSVALLMAVAEEVGLRSGQDLANQYKQGRVAVSLQIDMNNY---- 286  
 QY 373 LQFTGSEKARAIEMEWSLLQ--PLNITQV-----LSHGEG-----TDINFWI 413  
 Db 287 -----RGSADIVFITDYTDSNLTQFLTLLIDEXLPFLTYGYDRCGVACSDHASWH 337  
 QY 414 QAGVPGASLDDLYK-YEFFHSHGDTMTWDPQGMVAAVMAVSYV 462  
 Db 338 KAGSAAAMPFSKPKDYNPKLIHTSQDTLANSDPYGNHNAVFTKLGALAVYI 387

RESULT 13  
 F84215  
 antipeptidase [imported] - Halobacterium sp. NRC-1  
 C.Species: Halobacterium sp. NRC-1  
 C.Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C.Accession: F84215  
 R.Ng, W.V.; Kennedy, S.P.; Mahdizadeh, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.  
 ; Leitshauer, B.; Keller, K.; Cruz, R.; Danon, M.O.; Hough, D.W.; Maddocks, D.G.; Jablor  
 Jung, K.H.; Alam, M.; Frelais, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A.Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li,  
 A.Title: Genome sequence of Halobacterium species NRC-1.  
 A.Reference number: A84160; MUID:20504483; PMID:11016950  
 A.Accession: F84215  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-433 <STD>  
 A.Cross-references: GB:AE004437; NID:g10580170; PIDN:AAG19090.1; GSPDB:GN00138  
 C.Genetics:  
 A:Gene: ywad

Query Match 6.5%; Score 156; DB 2; Length 433;  
 Best Local Similarity 21.3%; Pred. No. 0.00043;  
 Matches 96; Conservative 72; Mismatches 167; Indels 116; Gaps 20;

QY 64 YERLALVDTVPRLSGSKULEKAIQIMYONLQODGLEKYLEPVRNIPMERGESAVML 123  
 Db 15 WTHLQRLVDT-PTWMASSDGERAAATRDALQHA-DDAWLDTFPVQGWTRGD----- 66  
 QY 124 EPRTHKIALILGSSIGTPREGITAEVLVWTSFDELQORASEARKIVVYNQPIYNSRT 183  
 Db 67 -----SAIETPAD-TTHADITALP-----RSPAGTAGERV-----D 96  
 QY 184 VQYRTQGAVEAAKV-GAL-----ASLRSVASFSIYSPH----- 216  
 Db 97 LSYGLPGDFEDADIDGAIWVAADVDPMDYRHLHRRKRYHAANAAGAFYVMHNVAGCL 156  
 QY 217 --TG-IOBYQDGVKIPACTIVEDAEMSGRMASHGIKIVIOQKMAKTY----- 263  
 Db 157 PATSVGTGDDPDIPIEIVAVGSTETGRLSRFFHD--TVTLVLDVADTDTQSONYAT 213  
 QY 264 --PDTSEFNTVAETGSKTPEQVVLVSGHLSWDVQGGAMDGGGAFISWEALSILKDLG 321  
 Db 214 LGPDTD-----AEI-----LVTSYDADHDITGANDNGGTM--AVLAAILA 255  
 QY 322 LRPK---RTLVLVMTAEQGVGAFOYOLHKVNIYSVLSMESDAGTFLPTGLQFTG- 377  
 Db 256 ARSALDTRVHFVCFGAEBVGLVGS--HHDAAGQDLDDVAVNLDDGVVRERTLKLVTHR 313  
 QY 378 ---SEKARAIEMEWSLLQPLNITQVLSHGEGTIDNFWIOAGVPGASLDDLYKFFEH 433  
 Db 314 FDALGDAAANAVSD--TFAPHTVAPALN--PHSDHWAAYVGVGVGHATSDAAGRGWG 368

Qy 434 HSHGDTMTMDPKQNNVAWVSVYVAD 464  
 Db 369 HTSADTLDDKLELRTFREHAIIILAAALAD 399

# RESULT 14

hypochemical protein XF0820 (imported) - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 15-Sep-2000

C:Accession: G82759

C:Title: The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequences

C:Accession: G82759

C:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: G82759

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-529 <SIM>

A:Cross-references: GB:A8003921; GB:A8003849; NID:g9105710; PIDN:AA83630.1; GSPDB:GN001

A:Experimental source: strain 9a5c

R:Stinson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carver, H.

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitching, J.P.; Krieger, J.B.; Kuramae, E.E.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, H

A:Authors: Martins, E.M.F.; Matubuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv

M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, W.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF0820

C:Superfamily: Xylella fastidiosa hypothetical protein XF0820

Query Match 6.2%; Score 150.5; DB 2; Length 529;

Best Local Similarity 22.4%; Pred. No. 0.0016;

Matches 111; Conservative 76; Mismatches 188; Indels 121; Gaps 25;

Qy 57 GKQNRSTVERLALVDTYGPRLSGSKLEKALQIMYQLODGL- -KHLERPRITHW 113

Db 4 GKRMSEVYKOLA-SDAFQGRAPSGAGEQKTIAVLEQFAAGLQPVGDVQADGRKL- -W 59

Qy 114 ERG- -EESAVMLEPRI- -HKIALIGLSSIGTPPEG- -ITAEVLVTSF- - 156

Db 60 TQAVPLRTTEILGPRVLLGAPGRKSLTQGOEIAVRAPLDASGAVLAQAPLVFVGV 119

Qy 157 - -DELORRASEARAKIV- -YNQP- -YINRSYVQYRTQAVENA 195

Db 120 KAPERHMDDF- -KGVDLKGLIAVMLINDPFTGQDFDGKMTYGGMTKFE- -EAA 174

Qy 196 KVGALASLI- - -RSVASFSIYSPHTGIGQVQDV- -PKIPACTTVEDEMS 243

Db 175 RQGLALIVHETAPASYGMAIVASSTNSMFDVRAPEAVHPKL- -TAMVQHDVAVOLF 233

Qy 244 RMAS- -HGKIVLQ- -LKGAKATYPD- -TSEFNVAEITTSKYPEQVYL 286

Db 234 KDAIGDFTTLIKKASRPFKVLLKQTLSDRYRKSEVITSHVAALQCAHDETVI 293

Qy 287 VSGHLSMDVG- - -QGAMDDGGAFTSWEALSLIKDL- -GLRPRKRLTLVLT 334

Db 294 YSAHMDHLGVGAPDAKGPITFNGALDNAG- -TALLLELRGAFRGPQPSVFLAVT 350

Qy 335 AEEQGVGAFOY- - -YQHKVINSYSLVWESDAGTFLPTGLQFTSGEAKAIMEEM 389

Db 351 ABEKGLDSEYASNPVLEK- -TVAIVIMDVMPQGPTRDGIYTAKLDDLMLKQVA 409

Qy 390 S- - -LLOPLNITQVLSHGEGDINFWIQAGVPGASLLDDYKYFFHHSHGDTMTMDPK 446

Db 410 AGWLRATLPTPEAGHFFRSDHFSFAKGIPLIS- - -YSAGQMEV- - 453

Qy 447 QMNVAAVWAVSVTV 462

Db 454 - -GVVAAGKAALEDYTV 468

# RESULT 15

JC7819 metallopeptidase II precursor - Alteromonas sp. (Strain O-7)

C:Species: Alteromonas sp. (Strain O-7)

C>Date: 03-Jun-2002 #sequence\_revision 03-Jun-2002 #text\_change 03-Jun-2002

C:Accession: JC7819

R:Miyamoto, K.; Tsujibo, H.; Nukui, E.; Itoh, H.; Kaitzu, Y.; Imanori, Y.

BioSci. Biotechnol. Biochem. 66, 416-421, 2002

A:Title: Isolation and characterization of the genes encoding two metalloproteases (Mpr1

A:Reference number: JC7818; PMID:11999419; MUID:21994062

A:Accession: JC7819

A:Molecule type: DNA

A:Residues: 1-609 <MTY>

A:Cross-references: DDBJ:AB063611

A:Experimental source: strain O-7

C:Comment: This enzyme, an extracellular alkaline metalloprotease, showing high similari

C:Genetics:

A:Gene: mpr11

Query Match 6.2%; Score 149.5; DB 2; Length 609;

Best Local Similarity 20.7%; Pred. No. 0.0023;

Matches 96; Conservative 70; Mismatches 158; Indels 139; Gaps 21;

Qy 14 LLSLCSKALCKNGISKRTFEETKEETIASGQVAKAIAINLAVYKQNRSEYERLALVDT 73

Db 8 LLSTIA- - -CLSG- - -FSEASEEYVWMTD- -AKSSQHPQVNSIRGTHFQSLT- - 52

Qy 74 VGPRLSGSKNLE- -KAIQIMYQNLQODGLEKHLERPVNI PMMERGEESAVMLEPRHKIAL 132

Db 53 - - - - -PNSVYKALKVP- -ANLHS- - -HL- - - -SARWDEYHRCG- 82

Qy 133 LGLGSSIGTPPEGITAEVLVTSFDELORRASE- -ARGKIVVYNOPYINYSRTVOYRTQGA 191

Db 83 - - - - -GFAH- - - - -DSLEQASVLAQESVNSPATLVNSIDNPTRVNSL 123

Qy 192 VEAQKVALSLRSVASF- -SIYSPHTGI- - -QEVDGVPKIPTACTIVEDAEM 242

Db 124 ISKVSTVNLDSVNSLTSFYNNRYTSQTGIDAAAVKQWSD- - - - -IAASRSDIS 174

Qy 243 SRMASHGKIKIVIQKMGAKTYPTDSEFNVAEITGSKYPEQVYLVSGHLSMD- - - - -V 296

Db 175 VEYISH- - - - -SMAOSSVITVTPGSELSEIYIIGHLDSINOSNSGL 218

Qy 297 GQGAMDDGGAFTSWEALSLIKDLGLRPRKRLTLVLTAE- - -OGGVGAFOYVQLHKVN 353

Db 219 APGADDNASGAVLVTALSAIVEDGYKPRRTIQMGFAAEVGLRGSKALAAQVSSQGN 278

Qy 354 I- - - - -SNYSLVWESDAGTFLPTGLQFTSGEAKAIMEEMSLDPLNTQV 400

Db 279 VVGKTPQDMTGRNRSADIVMTD- - - - -YTNASQNPQGLIETVPLSLAYVD 328

Qy 401 LSHGEGDINFWIQAGVPGA- - -SLDDLKYKFFHHSHGDT 439

Db 329 QCGVGCSDHASWYQGFPAAMPPEPSRSEINRKI- - -HTNNDT 368

Search completed: December 22, 2003, 11:52:59  
 Job time : 42.6646 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:34:20 ; Search time 29.9104 Seconds  
(without alignments)  
742.104 Million cell updates/sec

Title: US-09-745-763-36

Perfect score: 2417  
Sequence: 1 MKELIFAFGCVHLLSLCSG.....AVNAVSVYVADMEMLPRS 472

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	188	7.8	740 1	NID2_HUMAN
2	176	7.3	752 1	FOH1_RAT
3	175.5	7.3	776 1	TPR1_CHICK
4	174	7.2	752 1	FOH1_MOUSE
5	172.5	7.1	705 1	GCP2_ARATH
6	169	7.0	455 1	YRAD_BACSU
7	166	6.9	751 1	FOH1_PIG
8	165	6.8	750 1	FOH1_HUMAN
9	161.5	6.7	751 1	GCP2_CAEEL
10	153.5	6.4	769 1	TPR1_PELCA
11	147.5	6.1	504 1	AMPX_VIBBR
12	143.5	5.9	622 1	TPR1_RAT
13	143.5	5.9	740 1	NIDL_HUMAN
14	142	5.9	798 1	TPR2_MOUSE
15	138.5	5.7	284 1	APX_STRGR
16	137.5	5.7	537 1	APB3_YEAST
17	132	5.5	801 1	TPR2_HUMAN
18	130.5	5.4	770 1	TPR1_CANFA
19	129.5	5.4	811 1	YV96_YEAST
20	126	5.2	760 1	TPR1_HUMAN
21	125.5	5.2	413 1	YBS4_YEAST
22	123.5	5.1	745 1	NIDL_RAT
23	122	5.0	757 1	TPR1_CRICR
24	114	4.7	763 1	TPR1_MOUSE
25	113.5	4.7	794 1	YV99_SCHPO
26	110.5	4.6	456 1	SR54_THEAC
27	109.5	4.5	541 1	PTCD_MOUSE
28	105.5	4.4	470 1	PEPV_LACDL
29	105	4.3	411 1	ALIC_ECOLI
30	102	4.2	1332 1	XKDO_BACSU
31	101	4.2	490 1	RBL_ODOSI
32	100.5	4.2	933 1	VGLB_HSVAI
33	99	4.1	416 1	PGK_GLOMO

34	99	4.1	843 1	NUOG_STRCO	O9xar0 streptomyc
35	98.5	4.1	809 1	DCML_OLICA	P19919 oligotropha
36	98.5	4.1	1396 1	VCAP_VZVD	P09245 varicella-z
37	98.5	4.1	1485 1	TOP2_SCHPO	P08096 schizosach
38	97.5	4.0	145 1	CBPG_PSSS6	P06621 pseudomonas
39	97	4.0	886 1	YIH3_YEAST	P40511 saccharomyc
40	96.5	4.0	811 1	HSP7_YEAST	P33416 saccharomyc
41	96	4.0	385 1	SUCC_BACSU	P80886 bacillus su
42	96	4.0	693 1	GUAA_HUMAN	P49915 homo sapien
43	95.5	4.0	340 1	SRR_HUMAN	O9gzt4 homo sapien
44	95.5	4.0	541 1	PTCD_RAT	O88618 r formimido
45	95	3.9	368 1	FTZ1_ARCFU	O29715 archaeoglob

## ALIGNMENTS

RESULT 1  
NID2\_HUMAN STANDARD; PRT; 740 AA.  
AC O9Y3Q0;  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE N-acetylated-alpha-linked acidic dipeptidase II (EC 3.4.17.21)  
DE (NALADase II).  
GN NALAD2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N. A.  
RC TISSUE=Lung carcinoma;  
RX MEDLINE=99185063; PubMed=10085079;  
RA Pangalos M.N., Neels J.-M., Somers M., Verhasselt P., Bekkers M.,  
RA van der Helm L., Fraiponts E., Ashton D., Gordon R.D.;  
RT "Isolation and expression of novel human glutamate carboxypeptidases  
RT with N-acetylated alpha-linked acidic dipeptidase and dipeptidyl  
RT peptidase IV activity.";  
RT J. Biol. Chem. 274:8470-8483(1999).  
RL -1- FUNCTION: Has N-acetylated-alpha-linked-acidic dipeptidase  
CC (NALADase) activity. Also exhibits a dipeptidyl-peptidase IV type  
CC activity. In vitro, cleaves Gly-Pro-AMC.  
CC -1- CATALYTIC ACTIVITY: Release of an unsubstituted, C-terminal  
CC glutamyl residue, typically from Ac-Asp-Glu or pteroyl-gamma.  
CC -1- COFACTOR: Zinc; Binds two ions per subunit. Required for NALADase  
CC activity.  
CC -1- ENZYME REGULATION: Inhibited by quisqualate.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Potential).  
CC -1- TISSUE SPECIFICITY: Highest expression in the testis. Also found  
CC in ovary and spleen. Weak expression in prostate, heart and  
CC placenta. In brain, expressed in striatum, parietal cortex and  
CC ventral striatum with lower levels in hippocampus, brain stem,  
CC putamen and superior colliculus.  
CC -1- DOMAIN: The NALADase activity is found in the central region, the  
CC dipeptidyl peptidase IV type activity in the C-terminal.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AJ012370; CAB39967.1; -  
CC MEROPS: M28.012; -  
CC Genew: HGNC:14526; NALAD2.  
CC GO: GO:0016021; C:integral to membrane; NAS.  
CC GO: GO:0004180; F:carboxypeptidase activity; NAS.  
CC GO: GO:0016805; F:dipeptidase activity; NAS.

DR GO: GO:0004274; F: dipeptidyl-peptidase IV activity; NAS.  
 DR GO: GO:0006508; P: proteolysis and peptidolysis; NAS.  
 DR InterPro: IPR003137; PA.  
 DR Pfam: PF02225; PA; 1.  
 DR Pfam: PF04389; peptidase\_M28; 1.  
 DR Pfam: PF04253; TFR dimer; 1.  
 KW Hydroxylase; Carboxypeptidase; Metalloprotease; Zinc; Aminopeptidase;  
 KW Dipeptidase; Serine protease; Transmembrane; Glycoprotein;  
 KW Signal-anchor; Multifunctional enzyme.  
 FT DOMAIN 1 7  
 FT TRANSMEM 8 31  
 FT DOMAIN 32 740  
 FT DOMAIN 264 577  
 FT ACT\_SITE 414 414  
 FT ACT\_SITE 618 618  
 FT ACT\_SITE 656 656  
 FT ACT\_SITE 679 679  
 FT ACT\_SITE 367 367  
 FT METAL 377 377  
 FT METAL 415 415  
 FT METAL 443 443  
 FT METAL 543 543  
 FT CARBOHYD 111 111  
 FT CARBOHYD 143 143  
 FT CARBOHYD 185 185  
 FT CARBOHYD 314 314  
 FT CARBOHYD 449 449  
 FT CARBOHYD 603 603  
 FT CARBOHYD 628 628  
 SQ SEQUENCE 740 AA; 83591 MW; 040624D691ECF879 CRC64;

Query Match 7.8%; Score 188; DB 1; Length 740;  
 Best Local Similarity 20.0%; Pred. No. 6, 9e-06;  
 Matches 112; Conservative 85; Mismatches 186; Indels 176; Gaps 27;

DR KAON-----RSYERLALVDVTPRLSGSKN---LEKAIQIYQNIQDGLKHALEP-V 108  
 56 KAENKSFLESTFKL-----PHLAGTEQNFLLAKKIQTOMKRGFSDALVIVYDVL 107  
 DB 109 RIPIHMEGESAVMLEPRIKIALIIGLSSIGTPEGITAEVLVVTSDDELORRASEARG 168  
 108 SYPEETNANYISIVDE---HETEIFKT-SYLEPPDGYENVNIVPYNFAFGQM-DEG 162  
 QY 169 KIVVNOYINYSRVVQY-----RPGQVEAKKAL-----ASLIRYASPSI 212  
 DB 163 DLV-----YVNAKTEDPFKLEREMGINCTKIVARYGKIPRGKVKVNAMLAGIIL 217  
 QY 213 YS-----PHTGIQ-----EYQD 224  
 DB 218 YSDPADYFAPVQYPRPKGNLPGTAQKGNVNLNAGADPLTGYPAKEYTFRDLVEBGV 277  
 QY 225 GVPKIPTACTIVEDAEMMSRMASHGIKIVIQKMAKT-----YPTDSF----- 269  
 DB 278 GIPRIPIVHPIGVNDAEILRLRYLG-GIAPDKSKMKALNVYSIGFGFSDSFRKRVNHV 336  
 QY 270 -----NTVAITTSKYEQVIVVUSGHLDMDVCGANDDGGAFISME---ALSL 316  
 DB 337 YNINKITRIYVNVGTRGSVEPDRVYVILGGRHDSWVF-GALDPTSGAVVLOEILARSPGK 394  
 QY 317 IKDLGLRPKRTLRVLMTAEQGVAGFYQYOLHKVNISNYSLVNESDAGTFLPTGLPT 376  
 DB 335 LMSKMRRTIRITIPASWDAEERGLLGSTMAE-----ENVKILDERST-AIINDSSIE 447  
 QY 377 GSEKARA-----IMEVNSLLOPL-----NITVYLSHGE 405  
 DB 448 GNYTRRVCTPLLYLVLTKEIPSPDDGFESKSLYESMLEKDSPEKNLPRINKLGS 507  
 QY 406 GTDIFPMIO-AGVPGA-----SLDDLYKYFFHHSHDQNTTW-----DP---KOMVYA 451  
 DB 508 GSDFAFYQRLGIASGARARYTNKKTKDYSSYPVHTIYETELVEKFPYDFPKKOLSA 567  
 QY 452 AAVMAVSVVADMEEMLP 470

DB 568 QLRGALV-YELVD-SKILP 584  
 RESULT 2  
 FOHL RAT STANDARD; PRT; 752 AA.  
 AC P70627;  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Glutamate carboxypeptidase II (EC 3.4.17.21) (Membrane glutamate  
 DE carboxypeptidase) (MGCP) (N-acetylated-alpha-linked acidic dipeptidase  
 DE I) (NAAADase I) (Pteroylpolymy-gamma-glutamate carboxypeptidase)  
 DE (Folyl)polymy-gamma-glutamate carboxypeptidase (FGCP) (Folate hydrolase  
 DE 1) (Prostate-specific membrane antigen homolog).  
 GN FOHLI OR NAAAD1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxId=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Hippocampus;  
 RX MEDLINE=98041505; PubMed=9375657;  
 RA Bzdega T., Turi T., Wroblewska B., She D., Chung H.S., Kim H.,  
 RA Neale J.H.;  
 RT "Molecular cloning of a peptidase against N-acetylaspartylglutamate  
 RT from a rat hippocampal cDNA library."  
 RL J. Neurochem. 69:2270-2277 (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
 RX MEDLINE=98169524; PubMed=9501243;  
 RA Luchli-Carter R., Berger U.V., Barczak A.K., Emma M., Coyte J.T.;  
 RT "Isolation and expression of a rat brain cDNA encoding glutamate  
 RT carboxypeptidase II."  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:3215-3220 (1998).  
 RN [3]  
 RP SEQUENCE OF 284-752 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=96149377; PubMed=8570628;  
 RA Carter R.E., Feldman A.R., Coyte J.T.;  
 RT "Prostate-specific membrane antigen is a hydrolase with substrate and  
 RT pharmacologic characteristics of a neuropeptidase."  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:749-753 (1996).  
 RN [4]  
 RP SEQUENCE OF 284-752 FROM N.A.  
 RC TISSUE=Brain;  
 RA Luchli-Carter R., Berger U.V., Barczak A.K., Emma M., Coyte J.T.;  
 RL Submitted (DBS-1997) to the EMBL/Genbank/DBJ databases.  
 RN [5]  
 RP ALTERNATIVE SPLICING.  
 RA Bzdega T., She D., Turi T., Wroblewska B., Neale J.H.;  
 RT "Molecular cloning of alternatively spliced variants of the peptidase  
 RT against N-acetylaspartylglutamate (NAAAG) from human and rat nervous  
 RT systems."  
 RL Abstr. Soc. Neurosci. 24:579-579 (1998).  
 CC -1- FUNCTION: Has both folate hydrolase and N-acetylated-alpha-linked-  
 CC acidic dipeptidase (NAAADase) activity. Has a preference for tri-  
 CC alpha-glutamate peptides (By similarity). In the intestine,  
 CC required for the uptake of folate. In the brain, modulates  
 CC excitatory neurotransmission through the hydrolysis of the  
 CC neuropeptide, N-acetylaspartylglutamate (NAAAG), thereby releasing  
 CC glutamate.  
 CC -1- FUNCTION: Also exhibits a dipeptidyl-peptidase IV type activity  
 CC (By similarity). In vitro, cleaves Gly-Pro-AMC (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Release of an unsubstituted, C-terminal  
 CC glutamyl residue, typically from Ac-Asp-Glu or pteroyl-gamma-  
 CC -1- COFACTOR: Zinc; Binds two ions per subunit. Required for NAAADase  
 CC activity.  
 CC -1- ENZYME REGULATION: The NAAADase activity is inhibited by beta-  
 CC NAAAG, quisqualic acid and 2-(phosphonomethyl)glutaric acid (PMG).



CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Plasma membrane (By similarity).

CC -1- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing: Named isoforms=3;

CC Comment-Experimental confirmation may be lacking for some isoforms;

CC Name=1;

CC IsoId=P70627-1; Sequence=Displayed;

CC Name=2; Synonyms=Short form;

CC IsoId=P70627-2; Sequence=Not described;

CC Note=Probably inactive;

CC Name=3; Synonyms=Long form;

CC IsoId=P70627-3; Sequence=Not described;

CC -1- TISSUE SPECIFICITY: Widely expressed throughout brain regions with highest levels in the hippocampus, dentate gyrus, piriform cortex, choroid plexus of ventricles, pineal gland, anterior lobe of the pituitary gland and suprachiasmatic nucleus. High levels also found in the cerebral cortex, substantia nigra, pontine nucleus and the granule cell layer of cerebellum. Highly expressed in astrocytes and non-myelinating Schwann cells. Also expressed in kidney, localizing to the proximal brush border of the renal tube.

CC -1- DOMAIN: The NALADase activity is found in the central region, the dipeptidyl peptidase IV type activity in the C-terminal.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; U75973; AAC53423.1; -

DR EMBL; AF040256; AAC4067.1; -

DR EMBL; AF039707; AAB96759.1; -

DR MEROPS; M28.010; -

DR InterPro; IPR003137; PA.

DR Pfam; PF02225; PA; 1.

DR Pfam; PF04389; Peptidase\_M28; 1.

DR Pfam; PF04253; TFR\_dimer; 1.

KM Hydroxylase; Carboxypeptidase; Metalloprotease; Zinc; Aminopeptidase;

KM Dipeptidase; Serine protease; Transmembrane; Glycoprotein;

KM Signal-anchor; Multifunctional enzyme; Alternative splicing.

FT DOMAIN 1 19 CYTOPLASMIC (POTENTIAL).

FT TRANSHEM 20 44 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).

FT DOMAIN 45 752 EXTRACELLULAR (POTENTIAL).

FT DOMAIN 276 589 NALADASE.

FT ACT\_SITE 426 426 NUCLEOPHILE (NALADASE) (BY SIMILARITY).

FT ACT\_SITE 630 630 CHARGE RELAY SYSTEM (POTENTIAL).

FT ACT\_SITE 668 668 CHARGE RELAY SYSTEM (POTENTIAL).

FT ACT\_SITE 691 691 CHARGE RELAY SYSTEM (POTENTIAL).

FT METAL 379 379 ZINC 2 (BY SIMILARITY).

FT METAL 389 389 ZINC 1 AND 2 (BY SIMILARITY).

FT METAL 427 427 ZINC 1 (BY SIMILARITY).

FT METAL 455 455 ZINC 2 (BY SIMILARITY).

FT METAL 555 555 ZINC 1 (BY SIMILARITY).

FT CARBOHYD 48 48 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 78 78 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 123 123 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 155 155 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 197 197 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 338 338 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 461 461 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 478 478 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 615 615 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 640 640 N-LINKED (GLCNAC...) (POTENTIAL).

SEQUENCE 752 AA; 84539 MW; 5C0915A3B9C71E41 CR664;

Query Match 7.3%; Score 176; DB 1; Length 752;

Best Local Similarity 20.0%; Pred. No. 5,2e-05;

Matches 106; Conservative 78; Mismatches 172; Indels 174; Gaps 24;

QY 3 FLIAPFGVHLISLSCGAIKCKNGISKRTFEIEKEIASCGVAKAIIMLAVYKAQNR 62

DB 34 FIIQFLGWIITKSNDBSTSSVPGMKAFLOELKAE-----NIKRTLVNF----- 79

QY 63 SYERLALLVDVGRSLGSK-NLEKAIQIWMYQNLQODGLEKHLPEVRI-----PH 112

DB 80 -----TRTPHLAGTOHNFELAKOI-HAQWKEGGLDVELSDYDVLVSNKTHPN 128

QY 113 W-----ERGESAVMLEPRHKIAIILGLSSIGPP--EGITAEVLYVTSFDELGRASE 165

DB 129 YIILINDGNE-----IFKTSLAEL-----SPPEYENISDVPPYSAFSP--QGT 171

QY 166 ARGKIVVYNQPYINVSRT-----VOYRTQGAVAAPAKGAL-----ASLIRSVAS 209

DB 172 PEGDLV-----YNNVATDEPFKLERVMKINCSKVIYARIGVFRGNKYNQALQAKG 226

QY 210 FSIYSPT-----GIDRYOD----- 224

DB 227 IILVSDPADYFVPQKSYPPGMNLPGGGVORGNVLNNGAGDPLTPGYPANEVYRHEFT 286

QY 225 ---GVKIPFRACTVERAE-MMSRMASH-----GKIVI-----QLKM 258

DB 287 EAVGLPSIIPVHPPIGYDQAQLLEHMGSGAPDSSWKGLVPYNVGPFAGNFSKOKVXL 346

QY 259 GAKTYPD-TDSFNTVAEITGSKYPEQVYLVSGHLSMDVQGMDDGGAFFISWE--AL 314

DB 347 HHSYNNVTRIVYVIGLTKANVBDRIYLGHRDANVF--GGIDPQSGAAVHEIVRTF 404

QY 315 SLIDQLGRPKRTLRLVLTAEBOGVGAFQYQLHKNISNYSLVMSDAGTFLLPTGLQ 374

DB 405 GTLKCKGMRRRTILFASMDAEEFGILGSTEMABEH-----SRLQERGVAVINDSS 457

QY 375 FTSEKRAIMEEVMSLLOPLNITOVLSHREGDINWIOAGVPASLLD 424

DB 458 IEGNYTLRVDTPLMHS-LVYNLTKEPLSPD-----EGEGKSLYD 497

RESULT 3

TFRI CHICK STANDARD; PRT; 776 AA.

AC Q90997;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Transferrin receptor protein 1 (TFRI) (TR) (Tfr).

GN TFR.

OS Gallus gallus (Chicken).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

CC Gallus.

OX NCBI\_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Erythroblast, and Lymphoma;

FX MEDLINE=91340160; PubMed=1874449;

RA Gerhardt E.M., Chan L.-N.L., Jing S., Qi M., Trowbridge I.S.;

RT "The cDNA sequence and primary structure of the chicken transferrin receptor."

RL Gene 102:249-254 (1991).

CC -1- FUNCTION: Cellular uptake of iron occurs via receptor-mediated endocytosis of ligand-occupied transferrin receptor into specialized endosomes. Endosomal acidification leads to iron release. The apotransferrin-receptor complex is then recycled to the cell surface with a return to neutral pH and the concomitant loss of affinity of apotransferrin for its receptor. Transferrin receptor is necessary for development of erythrocytes and the nervous system (By similarity).

CC -1- SUBUNIT: Homodimer; disulfide-linked. Binds one transferrin molecule per polypeptide chain (By similarity).

CC -1- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).

CC -1- PTM: N- and O-glycosylated, phosphorylated and palmitoylated (By similarity).

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.  
 CC -1- CAUTION: A frameshift has been introduced after position 760 to  
 CC produce a C-terminal sequence homologous to other species  
 CC transmembrane receptor. The correct sequence is shown in fig.1 of  
 CC Ref.1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X55148; CA339035.1; ALT\_FRAME.  
 CC HSSP: P02786; ICX8.  
 CC InterPro: IPR003137; PA.  
 CC Pfam: PF04389; Peptidase M28; 1.  
 CC Pfam: PF04253; TFR dimer; 1.  
 CC Transmembrane; Glycoprotein; Receptor; Lipoprotein; Palmitate;  
 CC Signal-anchor; Endocytosis; Phosphorylation; Polymorphism.  
 CC DOMAIN 1 70  
 CC TRANSMEM 71 91  
 CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 CC (POTENTIAL).  
 CC EXTRACELLULAR (POTENTIAL).  
 CC DOMAIN 92 776  
 CC SITES 586 776  
 CC SITE 19 22  
 CC LIGAND-BINDING (BY SIMILARITY).  
 CC ENDOCYTOSIS SIGNAL.  
 CC CELL ATTACHMENT SITE (POTENTIAL).  
 CC PALMITATE (BY SIMILARITY).  
 CC MOD RES 23 23  
 CC DISTRES 101 101  
 CC CARBOHYD 261 261  
 CC CARBOHYD 326 326  
 CC CARBOHYD 331 391  
 CC CARBOHYD 738 738  
 CC VARIANT 581 581  
 CC VARIANT 736 736  
 CC SEQUENCE 776 AA; 85658 MW; DE293BAC41ED0623 CR64;  
 Query Match 7.3%; Score 175.5; DB 1; Length 776;  
 Best Local Similarity 21.8%; Pred. No. 5.9e-05;  
 Matches 83; Conservative 58; Mismatches 142; Indels 97; Gaps 18;  
 QY 136 GSSITPPEGITAELVLTSPDELGRASERAKIVNQPINNSRTYQRTQAVNA 195  
 DB 233 GSVSGKP---VYVNYGLKKDFEIIQKVASLNGTIVVAKGI---TLAEKVANAKKAG 285  
 QY 196 KVGALA---SLIRSVASFESI-----YSP-----HTGIGEOYO-DGVPKIPATC 233  
 DB 286 AAGLVMTYDSLMYGTITDILIPFGHAHLGDPRTYGFSPFNHTOPPPVSSGLPHIANQT 345  
 QY 234 ITVE-DAEMMSRM-----ASHGIKIV-----IQLMGA-KTYPPDTDSFNTVAE 274  
 DB 346 ISSAAARLRFKMDQDTCSEGWKGAHSCVTKTKQESQIMVXLDVNNMKRKLINIFGA 405  
 QY 275 ITGSKYPPQVVLVSHLDSMDVGGAMDGGCAFISWELSLIKL-----GLRPKTRRL 330  
 DB 406 IQGFEPFRYYVIGAQRDSS--GPGVAKAGTITALLBLARVTSIVNKEGKPPRSIIIF 463  
 QY 331 VLMTAEEOGVGAFQYVQLHKVNIYNSLVNEDAGTFLPTGLQGTSEKARAMEEYWS 390  
 DB 464 ASMSAGDVGAVGATW-----LEGYSAMLAHAKAFTYISLDAPVLGASHVYKISAPLTY 516  
 QY 391 LL-----OPTLITOVLSHGEGTDINFWIQAGVP-----GASLL----- 423  
 DB 517 MLIGSIMGVKVPAAVSSSLVRLGPD---WVKAVVPLGLDVAAPPLAYSGIPVLSGFG 573  
 QY 424 ---DLKYKFFPHSHGDTM 440  
 DB 574 YNKDEBYR---FLDTKGDTL 590

RESULT 4  
 FOH1 MOUSE  
 ID FOH1 MOUSE STANDARD; PRT; 752 AA.  
 AC O35409; Q9DCC2;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Glutamate carboxypeptidase II (BC 3.4.17.21) (Membrane glutamate  
 DE carboxypeptidase) (MGCP) (N-acetylated-alpha-linked acidic dipeptidase  
 DE I) (NALADase I) (Pteroylpolypoly-gamma-glutamate carboxypeptidase)  
 DE 1) (Prostate-specific membrane antigen (PSMA) (Foliate hydroxylase  
 GN FOH1 OR MOPSM OR NALAD1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 OX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NIH Swiss; TISSUE=Brain;  
 RX MEDLINE=21077532; PubMed=11210180;  
 RT "Cloning, expression, genomic localization, and enzymatic activities  
 RT of the mouse homolog of prostate-specific membrane antigen/NALADase/  
 RT foliate hydroxylase."  
 RL Mamm. Genome 12:117-123 (2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Iehi Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,  
 RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,  
 RA Schirni L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,  
 RA Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Baren G.,  
 RA Blake J., Botfield D., Boujarda N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., But C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Guenichon S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyshaw-Borle A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,  
 RA Hayashizaki Y.;  
 RL "Functional annotation of a full-length mouse cDNA collection."  
 CC Nature 409:685-690(2001).  
 CC -1- FUNCTION: Has both folate hydroxylase and N-acetylated-alpha-linked-  
 CC acidic dipeptidase (NALADase) activity. Has a preference for tri-  
 CC alpha-glutamate peptides (By similarity). In the intestine,  
 CC required for the uptake of folate. In the brain, modulates  
 CC excitatory neurotransmission through the hydrolysis of the  
 CC neuropeptide, N-acetylaspartylglutamate (NAAG), thereby releasing  
 CC glutamate.  
 CC -1- FUNCTION: Also exhibits a dipeptidyl-peptidase IV type activity  
 CC (By similarity). In vitro, cleaves Gly-Pro-MAC (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Release of an unsubstituted, C-terminal  
 CC glutamyl residue, typically from Ac-Asp-Glu or pteroyl-gamma-  
 CC -1- COFACTOR: zinc; Binds two ions per subunit. Required for NALADase  
 CC activity.  
 CC -1- ENZYME REGULATION: The NALADase and folate hydroxylase activities  
 CC are inhibited by quiescent acid.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Plasma membrane  
 CC (By similarity).  
 CC -1- TISSUE SPECIFICITY: Expressed predominantly in the hippocampal  
 CC region of the brain and in kidney. Lower levels in the ovary,  
 CC testis and mandibular gland.  
 CC -1- DOMAIN: The NALADase activity is found in the central region, the  
 CC dipeptidyl peptidase IV type activity in the C-terminal.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.

-1- CAUTION: There are amino acid differences between the sequence shown in fig.1 (Ref.1) and the sequence deposited in the database (AF026380). The sequence from fig.1 shows only 3 conflicts between Ref.1 and Ref.2. These are at AA positions 141, 240 and 287.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as the content is in no way CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

CC EMBL/ AF026380; AAB81971.1; -  
 DR EMBL/ AK002920; BAB22457.1; -  
 DR MEROPS: M28.010; -  
 DR MGD; MGI:1858193; Fc1h1.  
 DR GO; GO:0005887; C: integral to plasma membrane; IDA.  
 DR GO; GO:0004844; F: gamma-glutamyl hydrolase activity; IDA.  
 DR GO; GO:0004237; F: membrane dipeptidase activity; IDA.  
 DR GO; GO:0006760; P: folic acid and derivative metabolism; IDA.  
 DR InterPro: IPR003137; PA.  
 DR Pfam; PF02225; PA; 1.  
 DR Pfam; PF04389; Peptidase\_M28; 1.  
 DR Pfam; PF04253; TFR\_dimer; 1.  
 KM Hydrolyase; Carboxypeptidase; Metalloprotease; Zinc; Aminopeptidase;  
 KM Dipeptidase; Serine protease; Transmembrane; Glycoprotein;  
 KM Signal-anchor; Multifunctional enzyme  
 FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 23 44 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).  
 FT DOMAIN 45 752 EXTRACELLULAR (POTENTIAL).  
 FT ACT\_SITE 426 589 NUCLEOPHILE (NAILADASE) (BY SIMILARITY).  
 FT ACT\_SITE 630 630 CHARGE RELAY SYSTEM (POTENTIAL).  
 FT ACT\_SITE 668 668 CHARGE RELAY SYSTEM (POTENTIAL).  
 FT ACT\_SITE 691 691 CHARGE RELAY SYSTEM (POTENTIAL).  
 FT METAL 379 379 ZINC 2 (BY SIMILARITY).  
 FT METAL 389 389 ZINC 1 AND 2 (BY SIMILARITY).  
 FT METAL 427 427 ZINC 1 (BY SIMILARITY).  
 FT METAL 455 455 ZINC 2 (BY SIMILARITY).  
 FT METAL 555 555 ZINC 1 (BY SIMILARITY).  
 FT CONFLICT 141 141 F -> Y (IN REF. 2).  
 FT CONFLICT 178 178 F -> Y (IN REF. 2).  
 FT CONFLICT 219 219 V -> A (IN REF. 2).  
 FT CONFLICT 240 240 G -> A (IN REF. 2).  
 FT CONFLICT 287 287 E -> N (IN REF. 2).  
 FT CONFLICT 583 583 R -> G (IN REF. 2).  
 FT CONFLICT 625 625 E -> K (IN REF. 2).  
 FT CONFLICT 728 728 S -> N (IN REF. 2).  
 FT CONFLICT 749 749 M -> R (IN REF. 2).  
 SQ SEQUENCE 752 AA; 84635 MW; DE8F10946A50926A CRC64;

Query Match 7.2%; Score 174; DB 1; Length 752;  
 Best Local Similarity 19.9%; Pred. No. 7.2e-05;  
 Matches 105; Conservative 83; Mismatches 170; Indels 170; Gaps 24;

QY 3 FLTAFGVGHVLSGSKAICKNGISKRTPEIKETPIASCGDVAKAIINLATYGAQNR 62  
 DB 34 FLIGFLGWPFIKPNENATGNVSHSGMKPELHEIKAE-----NIKKFLYNP----- 79  
 QY 63 SYERLALLVDVGRGLSGSK-NLEKAIQIMYQNIQOGGLEKVNHEPRIPMERGESSAV 121  
 DB 80 -----TRTHLGLGTQNNFLAKOI-----HDQKEFGDLVELSLHYD-----V 117  
 QY 122 ML-----EPRIHKAIILGL-GSSI-----GTPREGITAEVLVVTSPDELQRRASEANGKI 170  
 DB 118 LLSVNTHTHPYISITINEDNEIFKTSLSQEPGPGYENISVVPYSAFSPQGT-PPGDL 176  
 QY 171 VVTVQPIYINSR-----VQRTQGAVEAKVGL--ASLIRSV-----ASFSTYS 214  
 DB 177 V-----FVNARTEDPFKLEREMKISCGKIVIRYGVFPGNNVKNVQLAGAKMTIYS 231

QY 215 PHT-----GIOEYOD-----GV 226  
 DB 232 DPADYFVPGVKGSPDGMNLPBGGVGNVNLNAGDPLTPGVPAHEHAYRHELTAVGL 291  
 QY 227 PKTPTACTIVEDAEMMSRNASH-----GIKI-----VIOLKGA 260  
 DB 292 PSLPVPPIGDDAQ-----KLEHMGCPAPDPSSWKGGLKPYNVNPGPAGNFSTOKYMMH 348  
 QY 261 KTYPD-TDSFNVAELTSGKYRPOVVLVSGHLSMDVVGQAMDDGAGFISWE---ALSL 316  
 DB 349 HSTKYTRIVNVIGTLGALPEPRVYILGGHRAWVF--GDIPOGAAVHHVNSFGT 406  
 QY 317 IKDLGLRPRKTLTLVMTAEEOGCGVAFQYQHLKXNISVNSVLMESDAGTFLPTGLQFT 376  
 DB 407 LKKGRRPRRTIIFASMDAEFFLLGSTWAEEH-----SLLDGKGVAYINADSSIE 459  
 QY 377 GSEKAPRIMEVMSLLQPLNTIYVLSHGECTDINFIQACVPASLID 424  
 DB 460 GNYTLRVDCPTPLMYSL-VVNLTKELQSPD-----EGFEKSKSLYD 497

RESULT 5  
 GCP2 ARATH STANDARD; PRT; 705 AA.  
 AC Q9M1S8;  
 DT 28-FEB-2003 (Ref. 41, Created)  
 DT 28-FEB-2003 (Ref. 41, Last sequence update)  
 DT 28-FEB-2003 (Ref. 41, Last annotation update)  
 DE Probable glutamate carboxypeptidase II (EC 3.4.17.21).  
 GN AMPI OR AT3654720 OR T5N23 80.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Landsberg erecta; TISSUE=slitques;  
 RX MEDLINE=21434423; Pubmed=11549767;  
 RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseid M.,  
 RA Fartmann B., Valle G., Bloecher H., Perez-Alonso M., Obermayer B.,  
 RA Delaeny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,  
 RA De Simone V., Choisne N., Artiguenave F., Robert C., Broetier P.,  
 RA Wincker P., Catolico L., Weissenbach J., Saurin W., Queller F.,  
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,  
 RA Wumbusch E., Drzonek H., Erfle H., Jordan N., Bangert P.,  
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,  
 RA Veezi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,  
 RA Conrad A., Hornischer K., Kauer G., Loehart T.-H., Nordstiek G.,  
 RA Reichelt J., Scharte M., Schoen O., Barynes M., Terol J., Climent J.,  
 RA Navarro P., Collado C., Perez-Perez A., Oetemaelder B., Duchemin D.,  
 RA Cooke R., Landie M., Berger-Llauró C., Punelle B., Masny D.,  
 RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,  
 RA Montfort A., Argitiro A., Flores M., Lisuori R., Vitale D.,  
 RA Mannhaupt G., Haase D., Schoof H., Rüd S., Zaccaria P., Kewes H.-W.,  
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,  
 RA Rooney T., Rizzo M., Walter A., Utzerback T., Fujii C.Y., Shea T.P.,  
 RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,  
 RA Pai G., Miltischer J., Sellers P., Gill J.E., Feldblum T.V.,  
 RA Preuss D., Lin X., Nieman W.C., Salzberg S.L., White O., Venter J.C.,  
 RA Frazer C.M., Kaneo T., Nakamura Y., Sato S., Kato T., Asamizu E.,  
 RA Sasamoto S., Kimura T., Ideasa K., Kawashima K., Kishida Y.,  
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A., Muraki A.,  
 RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,

QY	22	EYDGV-----	PKIPACTIVDAEWMASHG-----	249
Db	251	PGMVGVEGKSLDELVTRRFRFPJSLRRLAET--LALGGARAPLERNSGRV	308	
QY	250	-----IKIVQLKMGAKYPTDTSFNTVAETGSKYEQVVLVSGHLSWDVG	297	
Db	309	GPQRGPGRMVIINMTFOGEMMKIN-----NVVITIRSEADRVLIIGNHRDAMT--	361	
QY	298	QGAMDDGGAF-----ISWEASLIDGLRPRTLRVLVTAEEQGVGARFOYQLHKVN	353	
Db	362	YGAVDENSSTSLALLISRRFALLTKS-CMRPRRTILLCWPAEEFGMIGSTWIEERVLN	420	
QY	354	IS-----NYSLWESDAGTFLPRLGQFTGSEKARAIMBVMSLDQ-----	393	
Db	421	LGASAVAYLVNDCAVQG-SGFPAGATPQLDG-----LVLVLKLVDPDPAVGLTTEETF	473	
QY	394	--PLNTQVLSHGEGDITNFWIQAGVNPASLIDDLV--KYFFPHSHGDTMTVW----	445	
Db	474	KSQNNITIQLSLRVSDGDFSGFLHAGIPRI-----DMYTGADVPVYHTAFDSYDMMINHADP	529	
QY	446	K-QMNYA-AAWAVVSYVVAADMEMP 470		
Db	530	LFRHVAMAGIWLGLILLAD-EPLIP 555		
RESULT 6				
YWAD	BACSU	STANDARD;	PRT;	455 AA.
AC	P2152;			
DT	01-MAY-1992 (Rel. 22, Last sequence update)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Hypothetical peptidase ywad precursor (BC 3.4.11.-)			
GN	YWAD OR IPA-8R.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1423;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RX	MEDLINE=98044033; PubMed=9384377;			
RA	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,			
RA	Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,			
RA	Bourlles R., Bourster L., Brans A., Braum M., Brignell S.C., Bron S.,			
RA	Brocchi S., Bruschi C.V., Caldwell B., Capuano V., Catter N.M.,			
RA	Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,			
RA	Denizot F., Devine K.M., Dusterhoft A., Ehlich S.D., Emerson P.T.,			
RA	Enlin K.D., Errington J., Fabret C., Ferrari E., Foulger D.,			
RA	Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,			
RA	Ghm S.Y., Glaeser P., Goffeau A., Goldlighty E.J., Grandi G.,			
RA	Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,			
RA	Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,			
RA	Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,			
RA	Kobayashi Y., Koester P., Konigsegen G., Krogh S., Kumano M.,			
RA	Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,			
RA	Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigue C.,			
RA	Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,			
RA	Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,			
RA	Patro V., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,			
RA	Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,			
RA	Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadele Y.,			
RA	Sato T., Scanlan E., Schleich S., Schroter R., Scoffone F.,			

RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vaasart A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzneger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
 RT *subtilis*." ;  
 RT Nature 390:249-256(1997).  
 RL [3]  
 RN SEQUENCE OF 1-68 FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=92216127; PubMed=1806041;  
 RA Glaeser P., Kunst F., Debarbouille M., Vertes A., Danchin A.,  
 RA Dedonder R.;  
 RT "A gene encoding a tyrosine tRNA synthetase is located near sacs in  
 RT *Bacillus subtilis*." ;  
 RL DNA Seq. 1:251-261(1991).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X73124; CAAS1564.1; -;  
 DR EMBL: 299123; CAB15873.1; -;  
 DR EMBL: X52480; CA36725.1; -;  
 DR PIR: S39663; S39663.  
 DR MEROPS: M28. UPB-;  
 DR Subtilist; BGI0554; YWAD.  
 DR InterPro: IPR03137; PA.  
 DR Pfam: PF02225; PA; 1.  
 DR Pfam: PF04389; Peptidase M28; 1.  
 DR Hypothetical protein; Hydrolase; Aminopeptidase; Metalloprotease;  
 KW Zinc, Signal, Complete proteome.  
 FT SIGNAL 1 31  
 FT CHAIN 32 455  
 FT SEQUENCE 455 AA; 49450 MW; 89BE66EBB0CCB18 CRC64;  
 SQ  
 Query March 7.0%; Score 169; DB 1; Length 455;  
 Best Local Similarity 23.2%; Pred. No. 8.2e-05;  
 Matches 100; Conservative 83; Mismatches 174; Indels 74; Gaps 22;  
 QY 58 KAQRSYERLALVDVTPRLSGSKNLEKAIQIMYQNLQDGLKHLPEVPIPHMERGE 117  
 DB 44 KAKH-AISTISQLEAIGRIAGTAEEKSALLIASSRKLD-VKQGRNITDRLEGT 101  
 QY 118 ESAVMEPRHKAIALIGSSIGTPPEGITAEVIVTSTFDELQRRASARKIVVNOFY 177  
 DB 102 LSSAGRD-----ILQAAAGSAPTEBEGITAP-LYNAGLGXQDFTADAKKIMLSRGD 155  
 QY 178 INYERTQYRQGAWEAKVAKALASLI-RSAVASFIYSPHIGIOYQGVAKIPACTIV 236  
 DB 156 LTY-----YKAKNAEAA--GAKAVIYNNKESLVPMTPLNSGNV-----GIPVGIKK 203  
 QY 237 EDAMEMRMASHGIGIKIVQLKMGAKTYPDTSFNTVAITGSKYPE-----QVVLVSGHL 291  
 DB 204 EDGALVQOKKATYKL-----KAFNQNSQN-----ITGKKPKAKIHEDIYVTAHY 251  
 QY 292 DSWDVGGAMDGGGAPISWEALSLIKDLGAPKRTTALVLTAEQGVAFQY----- 347  
 DB 252 DSVFSGANDNGSGTSMLEMARVTK--SVPSDEIRFIKFAEGEILGSSHYDHL 309  
 QY 348 --QLHKVNSVSLVM-----ESDAGTFLP--GLQFGSKAKAIMEV-----MSLLQ 394  
 DB 310 EKEIKRSEV--NFNIDMGTSWEKASELYVNTLDGNSNVWSSSRPAEKIFDSLSLTQ 368  
 QY 395 LNTQVLSHGEGTD--INF-WIQAGVPGASLLDLKYKFFPHSHSDTWTVMDPKQN--- 449

DB 369 GSSDHFVPHAGIDISANFIW---GDPETEVEPWW-----HTPEDSIEHSKRLQOAG 419  
 QY 450 --VAAAVWAVV 458  
 DB 420 DLVTAAYEAV 430  
 RESULT 7  
 FOHL\_PIG STANDARD; PRT; 751 AA.  
 AC 077564;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Glutamate carboxypeptidase II (EC 3.4.17.21) (Membrane glutamate  
 DE carboxypeptidase) (MGC) (N-acetylated-alpha-linked acidic dipeptidase  
 DE I) (NALADase I) (Pteroylpolyl-gamma-glutamate carboxypeptidase)  
 DE (Folylpoly-gamma-glutamate carboxypeptidase) (FGCP) (Folate hydrolase  
 DE I) (Prostate-specific membrane antigen homolog).  
 GN FOHL OR NALADI.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 200-210 AND 471-483.  
 RC TISSUE=jejunal mucosa;  
 RX MEDLINE=98352082; PubMed=9685395;  
 RA Halsted C.H., Ling E.-H., Luthi-Carter R., Villanueva J.A.,  
 RA Gardner J.M., Coyle J.T.;  
 RT "Polyolpoly-gamma-glutamate carboxypeptidase from pig jejunum.  
 RT Molecular characterization and relation to glutamate carboxypeptidase  
 RT II." ;  
 RL J. Biol. Chem. 273:20417-20424 (1998).  
 RN [2]  
 RP ERRATUM.  
 RA Halsted C.H., Ling E.-H., Luthi-Carter R., Villanueva J.A.,  
 RA Gardner J.M., Coyle J.T.;  
 RL J. Biol. Chem. 275:30746-30746 (2000).  
 RN [3]  
 RP CHARACTERIZATION.  
 RX MEDLINE=86085936; PubMed=2867095;  
 RA Chandler C.J., Wang T.T., Halsted C.H.;  
 RT "Pteroylpolylglutamate hydrolase from human jejunal brush borders.  
 RT Purification and characterization." ;  
 RL J. Biol. Chem. 261:928-933 (1986).  
 CC -1- FUNCTION: Has both folate hydrolase and N-acetylated-alpha-linked-  
 CC acidic dipeptidase (NALADase) activity. Has a preference for tri-  
 CC alpha-glutamate peptides (By similarity). In the intestine,  
 CC required for the uptake of folate. In the brain, modulates  
 CC excitatory neurotransmission through the hydrolysis of the  
 CC neuropeptide, N-acetylaspartylglutamate (NAAG), thereby releasing  
 CC glutamate. Maximal activity at pH 6.0.  
 CC -1- FUNCTION: Also exhibits a dipeptidyl-peptidase IV type activity  
 CC (By similarity). In vitro, cleaves Gly-Pro-AMC (By similarity).  
 CC CATALYTIC ACTIVITY: Release of an unsubstituted, C-terminal  
 CC glutamyl residue, typically from Ac-Asp-Glu or Pteroyl-gamma-  
 CC glutamyl-zinc; Binds two ions per subunit. Required for NALADase  
 CC activity.  
 CC -1- ENZYME REGULATION: The NALADase activity is inhibited by  
 CC quinoxaline, ethanol, beta-NAG and 2-(phosphonomethyl) pentanedioic  
 CC acid (PMPA). Ethanol ingestion decreases the folate hydrolase  
 CC activity by 50%.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Plasma membrane  
 CC (By similarity).  
 CC -1- TISSUE SPECIFICITY: High expression in the duodenum and in the  
 CC jejunal brush-border membrane. Weak expression in kidney.  
 CC -1- DOMAIN: The NALADase activity is found in the central region, the  
 CC dipeptidyl peptidase IV type activity in the C-terminal.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration



RT acidic dipeptidase (NAALDase).";  
 RL J. Pharmacol. Exp. Ther. 286:1020-1025 (1998).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORM PSMA-1), AND CHARACTERIZATION.  
 RC TISSUE=Prostate;  
 RX MEDLINE=99185063; PubMed=10085079;  
 RA Pangalos M.N., Neels J.-M., Somers M., Verhasselt P., Bakkens M.,  
 van der Helm L., Fraiponts E., Ashton D., Gordon R.D.;  
 RT "Isolation and expression of novel human glutamate carboxypeptidases  
 with N-acetylated alpha-linked acidic dipeptidase and dipeptidyl  
 peptidase IV activity.";  
 RL J. Biol. Chem. 274:8470-8483 (1999).  
 RN [7]  
 RP SEQUENCE FROM N.A. (ISOFORMS PSMA-1 AND PSMA-2), AND VARIANT HHC  
 TYR-475.  
 RC TISSUE=Jejunum;  
 RX MEDLINE=20545101; PubMed=11092759;  
 RA Devlin A.M., Ling E.-H., Pearson J.M., Fernando S., Clarke R.,  
 Smith A.D., Halsted C.H.;  
 RT "Glutamate carboxypeptidase II: a polymorphism associated with lower  
 levels of serum folate and hyperhomocysteinemia.";  
 RL Hum. Mol. Genet. 9:2837-2844 (2000).  
 RN [8]  
 RP SEQUENCE FROM N.A. (ISOFORM PSMA-5).  
 RA Peace D.J., Zhang Y., Holt G., Ferrer K.T., Heller M., Sosman J.A.,  
 Xue B.H.;  
 RT "Identification of three novel splice variants of prostate-specific  
 membrane antigen.";  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 RN [9]  
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS PSMA-3 AND PSMA-4).  
 RA Lipold S.E., Criley S.C., Coffey D.S.;  
 RT "Alternative Splicing of the prostate-specific membrane antigen.";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [10]  
 RP SEQUENCE OF 60-74, AND SUBCELLULAR LOCATION.  
 RC TISSUE=Prostatic carcinoma;  
 RX MEDLINE=99025849; PubMed=9809977;  
 RA Grauer L.S., Lawler K.D., Marignac J.L., Kumar A., Geel A.S.,  
 Wolfert R.L.;  
 RT "Identification, purification, and subcellular localization of  
 prostate-specific membrane antigen PSM protein in the LNCaP prostatic  
 carcinoma cell line.";  
 RL Cancer Res. 58:4787-4789 (1998).  
 RN [11]  
 RP ALTERNATIVE SPLICING.  
 RA Bodega T., She D., Turi T., Wroblewska B., Neale J.H.;  
 RT "Molecular cloning of alternatively spliced variants of the peptidase  
 against N-acetylaspartylglutamate (NAAG) from human and rat nervous  
 systems.";  
 RL Abstr. Soc. Neurosci. 24:579-579 (1998).  
 RN [12]  
 RP CHARACTERIZATION.  
 RX MEDLINE=98288196; PubMed=9622670;  
 RA Luthi-Carter R., Barczak A.K., Speno H.D., Coyle J.T.;  
 RT "Hydrolysis of the neuropeptide N-acetylaspartylglutamate (NAAG) by  
 cloned human glutamate carboxypeptidase II.";  
 RL Brain Res. 795:341-348 (1998).  
 RN [13]  
 RP DOMAIN STRUCTURE.  
 RX MEDLINE=97330810; PubMed=9187245;  
 RA Rawlings N.D., Barrett A.J.;  
 RT "Structure of membrane glutamate carboxypeptidase.";  
 RL Biochim. Biophys. Acta 1339:247-252 (1997).  
 RN [14]  
 RP MUTAGENESIS.  
 RX MEDLINE=99102317; PubMed=9882712;  
 RA Speno H.S., Luthi-Carter R., Macias W.L., Valentine S.L.,  
 Joshi A.R.T., Coyle J.T.;  
 RT "Site-directed mutagenesis of predicted active site residues in  
 glutamate carboxypeptidase II.";  
 RL Mol. Pharmacol. 55:179-185 (1999).  
 CC -1- FUNCTION: Has both folate hydrolase and N-acetylated-alpha-linked-

CC acidic dipeptidase (NAALDase) activity. Has a preference for tri-  
 CC alpha-glutamate peptides. In the intestine, required for the  
 CC uptake of folate. In the brain, modulates excitatory  
 CC neurotransmission through the hydrolysis of the neuropeptide, N-  
 CC acetylaspartylglutamate (NAAG), thereby releasing glutamate. Stable  
 CC at pH greater than 6.5. Isoforms PSM-4 and PSM-5 would appear to  
 CC be physiologically irrelevant. Involved in prostate tumor  
 CC progression.  
 CC -1- FUNCTION: Also exhibits a dipeptidyl-peptidase IV type activity.  
 CC In vitro, cleaves Gly-Pro-AMC.  
 CC -1- CATALYTIC ACTIVITY: Release of an unsubstituted, C-terminal  
 CC glutamyl residue, typically from Ac-Asp-Glu or pteroyl-gamma.  
 CC -1- COFACTOR: Zinc; Binds two ions per subunit. Required for NAALDase  
 CC activity.  
 CC -1- ENZYME REGULATION: The NAALDase activity is inhibited by beta-  
 CC NAAG, quinic acid, 2-(phosphonomethyl) pentanedioic acid  
 CC (FMPA) and EDTA. Activated by cobalt.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Plasma membrane.  
 CC The PSMA' isoform is cytoplasmic.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=6;  
 CC Comment=Experimental confirmation may be lacking for some  
 CC isoforms;  
 CC Name=PSMA-1;  
 CC IsoId=Q04609-1; Sequence=Displayed;  
 CC Name=PSMA-2;  
 CC IsoId=Q04609-2; Sequence=VSP\_005341;  
 CC Name=PSMA-3;  
 CC IsoId=Q04609-3; Sequence=VSP\_005342;  
 CC Name=PSMA-4;  
 CC IsoId=Q04609-4; Sequence=VSP\_005339, VSP\_005340;  
 CC Name=PSMA-5;  
 CC IsoId=Q04609-5; Sequence=VSP\_005337, VSP\_005338;  
 CC Name=PSMA';  
 CC IsoId=Q04609-6; Sequence=VSP\_005336;  
 CC -1- TISSUE SPECIFICITY: Highly expressed in prostate epithelium. Also  
 CC expressed, in the small intestine, brain, kidney, liver, spleen,  
 CC colon and the capillary endothelium of a variety of tumors.  
 CC Expressed specifically in jejunum brush border membranes. In the  
 CC brain, highly expressed in the ventral striatum and brain stem. In  
 CC the prostate, the PSMA' cytosolic isoform is the most abundant  
 CC form in normal tissue, the membrane-bound PSMA-1 form in primary  
 CC prostate tumors. The PSMA-2 isoform also found in normal prostate  
 CC as well as in brain and liver.  
 CC -1- INDUCTION: In the prostate, up-regulated in response to androgen  
 CC deprivation.  
 CC -1- DOMAIN: The NAALDase activity is found in the central region, the  
 CC dipeptidyl peptidase IV type activity in the C-terminal.  
 CC -1- PTM: The first two amino acids at the N-terminus of isoform PSMA'  
 CC appear to be cleaved by limited proteolysis.  
 CC -1- PTM: The N-terminus is blocked.  
 CC -1- DISEASE: Defects in FOLH1 may be a cause of hyperhomocysteinemia  
 CC (HHCC), a condition associated with increased risk of  
 CC cardiovascular disease, neural tube defects, and cognitive  
 CC deficits.  
 CC -1- MISCELLANEOUS: PSMA is used as a diagnostic and prognostic  
 CC indicator of prostate cancer, and as a possible marker for various  
 CC neurological disorders such as schizophrenia, Alzheimer's disease  
 CC and Huntington's disease.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; W99487; AAA60209.1; -;  
 CC DR EMBL; S76978; AAB33750.2; -;  
 CC DR EMBL; AF007544; AAC83972.1; -;







Db 450 NISLHVDTVPLEHVAIEASQVENSPEKERSRKRLTYDWMKVPDPKAGVPIKRVPG 509  
 QY 405 EGTD---INFWIAGVPGASLDDLYKFFHHSHGT-----MTWMDPK 446  
 Db 510 GGSDDHAFPLN---AGVP-----VINFTKNTWDTYPLVHTMYETPPSNHLLDPTD 559  
 QY 447 QMNVAAAV---WAVSYVVADEEMLP 470  
 Db 560 NLSVHKAIGVWMAELAKTFND-DVILP 585

RESULT 10.  
 TFRL FELCA STANDARD; PRT; 769 AA.  
 AC Q9MYZ3;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Transferrin receptor protein 1 (TfR1) (TR) (TfR) (Tftr).  
 GN TRC.  
 OS Felis silvestris catus (Cat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 OX NCBI\_TaxID=9685;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver.  
 RX MEDLINE=21165299; PubMed=11264378;  
 RA Parker J.S.U., Murphy W.J., Wang D., O'Brien S.J., Parrish C.R.;  
 RT "Canine and feline parvoviruses can use human or feline transferrin  
 RT receptors to bind, enter, and infect cells.";  
 RL J. Virol. 75:3896-3902(2001).  
 CC -1- FUNCTION: Cellular uptake of iron occurs via receptor-mediated  
 CC endocytosis of ligand-occupied transferrin receptor into  
 CC specialized endosomes. Endosomal acidification leads to iron  
 CC release. The apotransferrin-receptor complex is then recycled to  
 CC the cell surface with a return to neutral pH and the concomitant  
 CC loss of affinity of apotransferrin for its receptor. Transferrin  
 CC receptor is necessary for development of erythrocytes and the  
 CC nervous system (By similarity).  
 CC -1- SUBUNIT: Homodimer; disulfide-linked. Binds one transferrin  
 CC molecule per polypeptide chain (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).  
 CC -1- PTM: N- and O-glycosylated, phosphorylated and palmitoylated (By  
 CC similarity).  
 CC -1- MISCELLANEOUS: Canine and feline parvoviruses bind human and  
 CC feline transferrin receptors and use these receptors to enter and  
 CC infect cells.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AF276984; AAF81908.1; -  
 DR HSBP; P02786; ICD8.  
 DR InterPro: IPR003137; PA.  
 DR Pfam: PF02225; PA; 1.  
 DR Pfam: PF04389; Peptidase\_M28; 1.  
 DR Pfam: PF04253; TFR\_dimer; 1.  
 KW Transmembrane; Glycoprotein; Receptor; Lipoprotein; Palmitate;  
 KW Signal-anchor; Endocytosis; Phosphorylation.  
 FT DOMAIN 1 70 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 71 91 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT DOMAIN 92 769 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 578 769 LIGAND-BINDING (BY SIMILARITY).  
 FT SITE 20 23 ENDOCYTOSIS SIGNAL.

FT SITE 61 64 STOP-TRANSFER SEQUENCE.  
 FT SITE 655 657 CELL ATTACHMENT SITE (POTENTIAL).  
 FT LIPID 70 70 PALMITATE (BY SIMILARITY).  
 FT MOD RES 24 24 PHOSPHORYLATION (BY SIMILARITY).  
 FT DISULFID 92 92 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 101 101 INTERCHAIN (BY SIMILARITY).  
 FT CARBOHYD 107 107 O-LINKED (GLNAC. . .) (BY SIMILARITY).  
 FT CARBOHYD 260 260 N-LINKED (GLCNAAC. . .) (BY SIMILARITY).  
 FT CARBOHYD 326 326 N-LINKED (GLCNAAC. . .) (BY SIMILARITY).  
 FT CARBOHYD 731 731 N-LINKED (GLCNAAC. . .) (POTENTIAL).  
 FT CARBOHYD 736 736 N-LINKED (GLCNAAC. . .) (BY SIMILARITY).  
 SQ SEQUENCE 769 AA; 86270 MW; 47EA4FCA18A46D7 CRC64;

Query Match 6.4%; Score 153.5; DB 1; Length 769;  
 Best Local Similarity 21.0%; Pred. No. 0.0023;  
 Matches 101; Conservative 65; Mismatches 177; Indels 139; Gaps 24;

QY 76 PRLSGSKNLEKAIQIYQNLQDGLERY---HLEPRIPHWERGESAVALBPRIRKIA 131  
 Db 163 PREAGSGKDSLAFFIENRPRELQLSKAMDEHVKVQV---KGSAS-----NSVT 210  
 QY 132 ILGLGSS---IGRPPEGITA-EVLYVT-----SPDELQRASEARKIIVYN 174  
 Db 211 IVGTNSGAVYLVSPEGYVAYSRAATVGRVHANFGTKDFENLN---SPVNGSLIVR 267  
 QY 175 QPVYINSGRTVQVRQGAVERAKVGALASLI-RSVAASFSI-----YSP 215  
 Db 268 AGKITPFAKV-----ANAEFRALGVLYMDQAKPTINAEIPFGHAHLTGDPYTP 320  
 QY 216 -----HTGIOEYO-DGVPIKTPACTIVEDAEMMSRWASHGIKIVIOKNGAKTYPDTDS 268  
 Db 321 GPFSPNTQPPPSGSSGLPMIPQGISRANAE-----KLFGMWEGDGPESAMETDS 370  
 QY 269 -----FNTVAELTSGSKPEDEVLYVSGHDSWDVQCGAMD 303  
 Db 371 SCRLETSRNNVVKLSVNNVLKEIRIFVFGVIGKFEEDHYVVGAGORDAM--GPGAKS 428  
 QY 304 GGAFFISWEALSLIKDL-----GLRPKRTLLVMTAEQGVGAFOYQO--LHKVNI SNY 357  
 Db 429 SVGTALLLELARIISDMVLKGGFKPSSIVPASKNSADPGAVGTEVLEGLSLHLKAF 488  
 QY 358 SLVWESDAGTFPLPTGLQFTSGEKARAIMBEVM-----SLIQPLN-ITVYLSHG-E 405  
 Db 489 TYI-NLKVAILGNSNFKVASPLLSLIEKVMQVKKHPVGGSLYRSNNINKYKESLD 547  
 QY 406 GTDINFTIQGVPCASLI---DDLKYKFFHHSHGDTMTWMDP---KQNTVAAYAVAV 457  
 Db 548 NAAFPFLVAGSIPAVSFCFCEDTDYPTL-----GTTMDVYEKLQKVPLINKKARAALAE 601  
 QY 458 VS 459  
 Db 602 VA 603

RESULT 11  
 AMPX\_VIBPR STANDARD; PRT; 504 AA.  
 AC Q01693;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 OS Bacterial leucyl aminopeptidase precursor (EC 3.4.11.10).  
 OC Vibrio proteolyticus (Aeromonas proteolytica).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=671;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 107-136 AND 233-405.  
 RC STRAIN=ATCC 15338;  
 RX MEDLINE=92329552; PubMed=1627651;  
 RA van Heeke G., Denslow S., Watkins J., Wilson K., Wagner F.;  
 RT "Cloning and nucleotide sequence of the vibrio proteolyticus  
 RT aminopeptidase gene.";

Query Match	Similarity	Score	DB 1;	Length	DB 2;	Length
Beet local	70; Conservative	47; Mismatches	110; Indels	77; Gaps	12	
192	VEAAKVGGLAALISVSASF--IYSPHTGICQ-----EYODGVKPIPTACTIVEDAEMM	242				
123	VDAQCI---TGTISLESFTRRFYTTTSGAQAQADWIAEWMQALSLPMAV-----	171				
243	SRMASHGIKIVIQLMGAKTYPDIDSFNTVAETGSKTPEQVVLVSGHLS-----W	294				
172	-KQVSHS-----GYNQKSVVMTITGSEAPDEWIVIGCHLDSTIGSHTNEQ	215				
295	DVCGGAMDDGGAFISWALSLIDKLGAPKRTLRVLVMTAEQGVG---AFQYQGLH	350				
216	SVAPADADGAGIAAVTVIRIVLSNNPQPKRSIAFMAYAAEVLRSQDILANDQYKSEG	275				
351	KVNISNYSIWMESDAGT-----FLPTGLOFTGSEKARAIMEEVMSILOPLNITQVLSHGE	405				
276	KNVVSLALDMLTNTYKGSADQVVFITDYDSNFTQVLTQLMDEYL-----PSLITVGF	326				
406	GT-----DINFWIOAGVPGASLLDDLYK-YFFHSHGDTTWVDP-----KQNVV	450				
327	DTCGVACSDHASMHNAGYPAAMPFSKFNQVNPRIHTTQDILANDPFGSHAKKFTQLGL	386				
451	AAAV 454					

Db 387 AYAI 390

RESULT 12

ID	TFR1_RAT	STANDARD;	PRT;	622 AA.
AC	Q99376;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Transferrin receptor protein 1 (TFR1) (TR) (TFR) (Tfr) (Fragment).			
GN	TFR1 OR TFR.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Testis;			
RC	MEDLINE=91125359; PubMed=2126342;			
RA	Roberts K.P., Griswold M.D.;			
RT	"Characterization of rat transferrin receptor cDNA: the regulation of			
RT	transferrin receptor mRNA in testes and in Sertoli cells in culture.";			
RL	Mol. Endocrinol. 4:531-542(1990).			
CC	-1- FUNCTION: Cellular uptake of iron occurs via receptor-mediated			
CC	endocytosis of ligand-occupied transferrin receptor into			
CC	specialized endosomes. Endosomal acidification leads to iron			
CC	release. The apotransferrin-receptor complex is then recycled to			
CC	the cell surface with a return to neutral pH and the concomitant			
CC	loss of affinity of apotransferrin for its receptor. Transferrin			
CC	receptor is necessary for development of erythrocytes and the			
CC	nervous system (By similarity).			
CC	-1- SUBUNIT: Homodimer; disulfide-linked. Binds one transferrin			
CC	molecule per polypeptide chain (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).			
CC	-1- TISSUE SPECIFICITY: In testis, expressed in Sertoli cells,			
CC	peritubular myoid cells and in germinal cells. Highest levels in			
CC	Sertoli cells.			
CC	-1- PPM: N- and O-glycosylated, phosphorylated and palmitoylated (By			
CC	similarity).			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; M58040; AAA4273.1; -.			
DR	PIR; A34549; A34549.			
DR	HSSP; P02786; ICB8.			
DR	InterPro; IPR003137; PA.			
DR	Pfam; PF02225; PA; 1.			
DR	Pfam; PF04389; Peptidase_M28; 1.			
DR	Pfam; PF04253; TFR_dimer; 1.			
KW	Transmembrane; Glycoprotein; Receptor; Lipoprotein; Palmitate;			
KW	Signal-anchor; Endocytosis; Phosphorylation.			
FT	DOMAIN 1			
FT	NON TER 1			
FT	SITE 431 622			
FT	CARBONYD 109 510			
FT	CARBONYD 113 109			
FT	CARBONYD 113 113			
FT	CARBONYD 179 179			
FT	CARBONYD 584 584			
FT	CARBONYD 589 589			
SO	SEQUENCE 622 AA; 70152 MW; 831E4FC1DEB55703 CMC64;			

Query Match 5.9%; Score 143.5; DB 1; Length 622;

Best Local Similarity 19.4%; Pred. No. 0.0088;

Matches 90; Conservative 61; Mismatches 158; Indels 155; Gaps 21;

Qy 72 DTGCPRLSGS---KNLEKAIQIMYQNLQDGLKRVHLEPVRIIPHWERGESAVMLEPR-- 126

Db 13 NTYTPRAGSQKXENLAVYIENLFPHDK-----FSKV-----WR--DEHYKIQVKN 58

Qy 127 -IKKIALIGGSSIG--TPREGITAEVLVYTSDELORASBARGLVYVNOPIYNSRT 183

Db 59 VSONLVITINGSNIDPVEAPEGYVA-----FSKAGEVYGLVHANFGTKDPFE 107

Qy 184 VOYRTGAVAAKAGAL--ASLIRSVASFSI----- 212

Db 108 LANSVNGSLVIYRAGKTTPEKRYANASFAVGLYIMDKNTPTPVVEADLQFGHAHGT 167

Qy 213 ---YSP-----HTGQIEYO-DGVPKIPTACTIVEDAEMMSR----- 244

Db 168 GDPYTPPEPFPNTHQPPSSGSLPSIVGTISRAPKLFKFMENGCPSWNIDSSCKL 227

Qy 245 -MAHSGIKIVIQKMGAKTYPTDSEFNVAETGSKYRPEQVVLVSGHLDSDVQCAMD 302

Db 228 ELSONQNVKLTVN-----NVLKETRLINIFGVIKGYEPEPRYIVGAQRDAMGPGVAKSS 282

Qy 303 DGGAPFISWEAL--SLIKDLGLRPPKRLRLVMTAEQGGVG-----AFQ 345

Db 283 VGTGLLKLAAQVPSDMISKDGFPRRSITFASWTAGYGAVGPTWLEGLSLHLKAPT 342

Qy 346 YVQLHKV--NISVSLVMSDAGTFLPTGLQFTGSEKARAIMSEV----- 388

Db 343 YINDKVLVLTSPKV-----SASPLVYTL-----GIMQDVHPRIDGKLYRNSW 390

Qy 389 MSLLQPLINTQVLSHGEGTDINFIQAGVPGASLL---DDLKYK 429

Db 391 ISKIEELSL-----DNAAPFLAVSGIPAVSFCEDEDDYK 427

RESULT 13

ID	NDDL_HUMAN	STANDARD;	PRT;	740 AA.
AC	Q9VQQL; O43176;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	N-acetylated-alpha-linked acidic dipeptidyl-like protein			
DE	(EC 3.4.17.21) (NALADase L) (ileal dipeptidyl-peptidase) (100 kDa			
DE	ileum brush border membrane protein) (1100).			
GN	NALADL OR NALADASEL.			
OS	Homo sapiens (Human).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
CC	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5; 6; 7 AND 8).			
RC	TISSUE=Small intestine;			
RK	MEDLINE=99185063; PubMed=10085079;			
RA	Pangalos M.N., Neels J.-M., Somers M., Verhasselt P., Bekkers M.,			
RA	van der Helm L., Fraiponts E., Ashton D., Gordon R.D.;			
RT	"Isolation and expression of novel human glutamate carboxypeptidases			
RT	with N-acetylated alpha-linked acidic dipeptidase and dipeptidyl			
RT	peptidase IV activity.";			
RL	J. Biol. Chem. 274:8470-8483(1999).			
RN	[2]			
RP	SEQUENCE OF 419-740 FROM N.A.			
RC	TISSUE=Ileum;			
RK	MEDLINE=98049571; PubMed=9388249;			
RA	Schneider B.L., Thevananthar S., Moyer M.S., Walters H.C., Rinaldo P.,			
RA	Devarajan P., Sun A.Q., Dawson P.A., Ananthanarayanan M.;			
RT	"Cloning and characterization of a novel peptidase from rat and human			
RT	ileum.";			
RL	J. Biol. Chem. 272:31006-31015(1997).			
CC	-1- FUNCTION: NALADase-like activity unknown. Has no NAAG hydrolyzing			
CC	activity. Exhibits a dipeptidyl-peptidase IV type activity. In			
CC	vitro, cleaves Gly-Pro-Amc (By similarity).			
CC	-1- CATALYTIC ACTIVITY: Release of an unsubstituted, C-terminal			
CC	glutamy1 residue, typically from Ac-Asp-Glu or pteroyl-gamma-			
CC	-1- COFACTOR: zinc; Binds two ions per subunit (By similarity).			

FT			/FTID=VSP_005345. PSLSDGAGSDDPAPVFHFLGI -> PRLQDPSGCCPDSSGC DSPAQ (in isoform 5). /FTID=VSP_005346. Missing (in isoform 5). /FTID=VSP_005347. PSLSGLGCSODXAPVHFGLISSMDIAYTYDRSKTSARIYP T-> FRARLDQPSPTTPLOPLTRPLTMWISFWIRASAKIRLM PGQRGV (in isoform 6). /FTID=VSP_005348. Missing (in isoform 6). /FTID=VSP_005349. FSSHQAVARTASIVLR -> FEEDGDKHPETRICEARD (in isoform 7). /FTID=VSP_005350. Missing (in isoform 7). /FTID=VSP_005351. GPLVAIVKFKFEAAAL -> GMSPPDEPVWALNHHND (in isoform 8). /FTID=VSP_005352. Missing (in isoform 8). /FTID=VSP_005353. LIGST -> SPGLQ (IN REF. 2). E255AAC9EB05CA14 CRC64;
FT	VARSPLIC	502	
FT	VARSPLIC	523	740
FT	VARSPLIC	502	543
FT	VARSPLIC	544	740
FT	VARSPLIC	562	579
FT	VARSPLIC	580	740
FT	VARSPLIC	619	635
FT	VARSPLIC	636	740
FT	CONFLICT	420	424
SEQ	SEQUENCE	740 AA; 80620 MW; E255AAC9EB05CA14 CRC64;	
Query Match			
Beat Local Similarity 20.2%; Pred. No. 0.011; Length 740;			
Matches 94; Conservative 63; Mismatches 157; Indels 151; Gaps 21			
OY			
Db	68	PRLSGSKULEKAIOIMYONLC--ODGLEKHMLEPRRI---PHMERGESAMVLEPRRIK	129
OY			
Db	130	IATIGLAGSSI:::---GRP-----PEGITAEVLVYTS-----PDE	158
Db	117	VDIVGTTCGITIHSCRTTEENTVEGGGDVQVPYAAYAPSGTPOGLLYVANNGABEDFRE	176
OY			
Db	159	LORPASEARGKIIVYNOPYINSRTVOYRTQGAVEAAYKVALASLI RS-----	206
Db	177	LQTGCIKLEGITALTRYGVGVC-----RGAKAVAANKHGAVGLVYTDPADINDGLSSP	229
OY			
Db	207	--VASFSIYSHTTGQ----EY-----QDGYPKIPACTIIVE	237
OY			
Db	230	DEFEPNSMYLPDLGVGERGSYYIEFGDPLPYLPAPVSSFRVDLANVSGFPPIPTGIFQ	289
OY			
Db	238	DA-EMMSRM-----ASHGIKIVILQMKGAKTYPPTD-----SFMT	271
OY			
Db	290	DARDLLCTUNGLTAPATIQGALCGHYRRGPERPCDDPADSQVNAVSVNYRELNNSSNV	349
OY			
Db	272	VASITGSKYPEOVYLVSCHLDSWDVGOGAMDGGAFTSWEAL-----SLIKDLGRPKET	327
OY			
Db	350	LGIIRGAVEPRRYLVYGNHRDSM--VHGADVDPSSGTAIVLELSRVLGTLTKKGTRWPRRS	407
OY			
Db	328	LRVLVMTAEBOGCGCAFQYIQ--LHKVNINSYSLVMESDAGFLPTGLOFTSSSEKARA IM	385
OY			
Db	408	IVFASWGABEGELIGSTTEEFEEFFKKLOBERTAYVI-NVDISYEANATLRVQGTTPVOYV-	465
OY			
Db	386	EENWSILQPLNITOVLSHGEGEDINFWIQAGVPGASILLDDLKYKF	430
OY			
Db	466	--VESATK-----ETRSPGPG-----DLSTIDNMWIRTF	491
RESULT 14			
TFR2_MOUSE			
ID	TFR2_MOUSE	STANDARD.	PRT: 798 AA.
AC	O9UKX3; Q92016; Q9SMO9; Q9CPT2;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Transferrin receptor protein 2 (TFR2).		
GN	TFR2 OR TFR2.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		



FT CONFLICT 248 248 S -> L (IN REF. 2).  
 FT CONFLICT 287 287 A -> V (IN REF. 2).  
 FT CONFLICT 595 595 K -> E (IN REF. 1).  
 SQ SEQUENCE 798 AA; 88402 MW; FA6161FE3FFFAA4 CRC64;

Query Match 5.9%; Score 142; DB 1; Length 798;  
 Best Local Similarity 18.4%; Pred. No. 0.016;  
 Matches 104; Conservative 69; Mismatches 190; Indels 202; Gaps 22;

QY 9 FGGHLLSCSGKAIKCKGKISKEEKEEKEEASGCD--VAKAIIN----- 52  
 DB 91 FTGAFLGYYVAFVFGSCQ-----ACGDSVLVDEEDVNEDESGRTLLYMSD 134  
 QY 53 -----LAVYGAKNQSYERIALLVTVGPRLSGKNLEKAIWQNTQQDLEKY----- 103  
 DB 135 LQAMLRRLGEGRMEDTIRLTLR-----RVAGSKAMATLVQDIDKLSRQLDHWVWDT 190  
 QY 104 HLEPVRI-----HWERG-----ESSAVMLEPRIH-----KIALIGSSIG 140  
 DB 191 HVGGLQFPDPAHANTLHWVDADGVSQEQPLEDEPVYCPYSATGNATKGLVYAHYGRS-- 248  
 QY 141 TPPEGITAEVLVYVTFDELQRRASARAKIYVYNOPIYNSKTYQRTQGAVERAKVGL 200  
 DB 249 -----EDLODLKAKGVELAGSLHLVRVGTISFAQKV-----AVAQDFGAQ 288  
 QY 201 ASLIRSVASFISYSPH-TGIOEYQ-----DGVPK 228  
 DB 289 GVLIVPDSDFSQDPKPKGLSHQVGHVHLGTDDPTTPGPPSPNQTOPPEVSSGLPS 348  
 QY 229 IPTACTIVEDAEMMSRMAS-----HGKIKYIQKMGAKYTPD-----TD 267  
 DB 349 IPAQRTISADIDQLKRLTGPVAPQEWKGLSGSPYRLGPG--PDLRLVNNHNVSTP 404  
 QY 268 SPNTVAETGSKYEPQVVLVSGHLSWDVGGCAMPDGGAFISWALSILDL--GLRP 324  
 DB 405 ISNTFACIEGAEPHYYVIGAORDAM--GPGAAASAVGTILFLVTFSSMVSNGFRP 462  
 QY 325 KRTLRVMTAEBOGVGAFOYQLHKVNISYLSWMSDAGTFLPTGLQFTGSEKANA- 383  
 DB 463 RSLFLFISWDGDFSGSVGATEW-----LEGYLSVLHLKAVYVSLNSVLGDKKFLAK 515  
 QY 384 -----IMEEVM-----LLOPLNITQVLSHGEGTDIN 410  
 DB 516 TSPILVSLIENILKQVDSPNHSGQTLVQGVALTHTPSWABAVIQPLPM-----DSSAYS 568  
 QY 411 FWIQGVPGA--SLDDLYKTFEPH 433  
 DB 569 FTAFAGVPAVEFSFMEDDRVPFLH 593

RESULT 15  
 APX\_STRGR STANDARD; PRT; 284 AA.  
 ID APX\_STRGR STANDARD; PRT; 284 AA.  
 AC P80561;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Aminopeptidase (EC 3.4.11.-) (SGAP).  
 OS Streptomyces griseus.  
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1911;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=96270734; PubMed=8665903;  
 RA Maras B., Greenblatt H.M., Shoham G., Spungin-Bialik A., Blumberg S.,  
 RA Barra D.;  
 RT "Aminopeptidase from Streptomyces griseus: primary structure and  
 RT comparison with other zinc-containing aminopeptidases.";  
 RL Eur. J. Biochem. 236:843-846 (1996).  
 RN [2]  
 RP SEQUENCE OF 1-6.  
 RX MEDLINE=89338422; PubMed=2503378;

RA Spungin A., Blumberg S.;  
 RT "Streptomyces griseus aminopeptidase is a calcium-activated zinc  
 RT metalloprotein. Purification and properties of the enzyme.";  
 RL Eur. J. Biochem. 185:471-477 (1989).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).  
 RX MEDLINE=97201142; PubMed=9048953;  
 RA Greenblatt H.M., Almog O., Maras B., Spungin-Bialik A., Barra D.,  
 RA Blumberg S., Shoham G.;  
 RT "Streptomyces griseus aminopeptidase: X-ray crystallographic  
 RT structure at 1.75-A resolution.";  
 RL J. Mol. Biol. 265:620-636 (1997).  
 RN [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE=93185612; PubMed=8444149;  
 RA Ben-Meir D., Spungin A., Ashkenazi R., Blumberg S.;  
 RT "Specificity of Streptomyces griseus aminopeptidase and modulation of  
 RT activity by divalent metal ion binding and substitution.";  
 RL Eur. J. Biochem. 212:107-112 (1993).  
 CC -1- FUNCTION: SPECIFIC FOR LARGER HYDROPHOBIC ACIDS, ESPECIALLY  
 CC LEUCINE. NO CLEAVAGE OCCURS IF THE NEXT RESIDUE IS PROLINE.  
 CC -1- COFACTOR: BINDS 2 ZINC IONS.  
 CC -1- ENZYME REGULATION: BINDS A CALCIUM ION WHICH MODULATES THE  
 CC ACTIVITY OF THE ENZYME.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- MASS SPECTROMETRY: MW=29728; MW ERR=1.0; METHOD=Electrospray.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.  
 DR PIR; S66427; S66427.  
 DR PDB; 1XJO; 01-APR-97.  
 DR PDB; 1CP7; 03-MAY-00.  
 DR PDB; 1P20; 22-AUG-01.  
 DR PDB; 1P23; 22-AUG-01.  
 DR PDB; 1QO3; 03-MAY-00.  
 DR MEROPS; M28.003; -.  
 DR Pfam; PF04389; Peptidase\_M28; 1.  
 DR Hydrolase; Aminopeptidase; Zymogen; Zinc; Calcium; 3D-structure.  
 FT METAL 85 97 ZINC 1.  
 FT METAL 132 132 ZINC 2.  
 FT METAL 160 160 ZINC 1.  
 FT METAL 247 247 ZINC 2.  
 FT DISULFID 245 250  
 FT HELIX 6 22  
 FT TURN 23 25  
 FT STRAND 28 28  
 FT TURN 29 30  
 FT HELIX 32 47  
 FT TURN 48 49  
 FT STRAND 51 59  
 FT TURN 60 61  
 FT STRAND 62 71  
 FT STRAND 75 85  
 FT STRAND 88 88  
 FT TURN 90 91  
 FT TURN 95 98  
 FT HELIX 99 115  
 FT TURN 116 116  
 FT STRAND 121 128  
 FT HELIX 131 133  
 FT TURN 134 134  
 FT HELIX 136 144  
 FT HELIX 147 150  
 FT TURN 151 152  
 FT STRAND 153 159  
 FT STRAND 170 171  
 FT HELIX 176 189  
 FT TURN 190 190  
 FT STRAND 194 194  
 FT TURN 204 204  
 FT HELIX 205 210  
 FT TURN 211 212  
 FT STRAND 215 219

FT STRAND 225 225  
FT HELIX 228 234  
FT TURN 238 239  
FT STRAND 241 241  
FT TURN 244 247  
FT TURN 249 250  
FT TURN 253 254  
FT HELIX 258 275  
FT TURN 276 276  
SQ SEQUENCE 284 AA; 29722 MW; 67F1B80F8CASC4CC CRC64;

Query Match 5.7%; Score 138.5; DB 1; Length 284;  
Best Local Similarity 21.6%; Pred. No. 0.0067;  
Matches 63; Conservative 42; Mismatches 113; Indels 73; Gaps 11;

QY 227 PKIPTACITVEDAMGSRMASHGIKIV-----IQKMGAKTYPDT----- 266  
DB 2 PDIPLANVKAHLTQLSTIAANNNGNRAHGRPGYKASVDYVAKLDAAGYTTTLOQFTSGG 61  
QY 267 -DSFNTVAETGSKYPCQVVLVSGHLSMDVGGAMDGGGAFISWEALSLIKDIGLRPK 325  
DB 62 ATGYNLIANWPGD-PNKVLMAGAHLSVSSGAGINDNGSSAAVLETALAVSRAGYQPD 120  
QY 326 RTLRVLVMTAEEOGGVGA-FQYQOLHKVNIS-----NYSLVMSDAGTFLPTGLQFTGS 378  
DB 121 KHLRFAMWGAHELGLIGSKFYVNNLPSADRSKLAGYLNFDWIGSPNPGYFV-----Y 172  
QY 379 EKAAIMEEWSLQPLNI-TQVLSHGG-TDINFMIQAGVPGASL----- 422  
DB 173 DDDPVIKTFKNYPAGLNVPTIEIEGDSRSDHAFKKNVGVVGGGLFTGAGYTKSAQAQ 232  
QY 423 -----LDDLKYFFFFHSHGPTM-----TWMDPKQMNVAANAVALVS 459  
DB 233 KMGGTAGQAFPRCY-----HSCDSLSNINDTALDRNSDAAAHAHWTLS 277

Search completed: December 22, 2003, 11:50:00  
Job time : 32.9104 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:37:00 ; Search time 71.493 Seconds  
(without alignments)  
1703.674 Million cell updates/sec

Title: US-09-745-763-36

Perfect score: 2417  
Sequence: 1 MKFLIFAFGCVHLLSLCSG.....AWMAVSYYVADMEEMLP RS 472

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2417	100.0	472	4 Q9Y646	Q9Y646 homo sapien
2	2402	99.4	472	4 Q8NBZ1	Q8NBZ1 homo sapien
3	2285	94.5	541	1 Q9Y5X6	Q9Y5X6 homo sapien
4	2133	88.2	470	11 Q9WVJ3	Q9WVJ3 mus musculus
5	2117	87.6	472	11 Q9JLV0	Q9JLV0 rattus norv
6	2116	87.5	472	11 Q9Z1Y1	Q9Z1Y1 rattus norv
7	1907	78.9	433	11 Q70216	Q70216 mus musculu
8	759.5	31.4	493	11 Q76552	Q76552 acanthochei
9	701	29.0	140	4 Q9UNM8	Q9UNM8 homo sapien
10	655	27.1	414	16 Q8EBH7	Q8EBH7 shewanella
11	640	26.5	467	16 Q9A5A8	Q9A5A8 caulobacter
12	623	25.8	139	11 Q9R114	Q9R114 mus musculu
13	499.5	20.7	472	16 Q8PHB8	Q8PHB8 xanthomonas
14	420	17.4	375	16 Q8P625	Q8P625 xanthomonas
15	193.5	8.0	449	2 Q9JEU5	Q9JEU5 bacillus li
16	182	7.5	536	16 Q9H2Q8	Q9H2Q8 pseudomonas

17	176.5	7.3	571	16 Q9A4J1	Q9A4J1 caulobacter
18	169	7.0	393	2 Q82996	Q82996 aeromonas p
19	165	6.8	466	16 Q9K671	Q9K671 bacillus ha
20	165	6.8	719	4 Q8TAY3	Q8TAY3 homo sapien
21	163	6.7	683	16 Q8PGK5	Q8PGK5 xanthomonas
22	163	6.7	1483	5 Q93332	Q93332 caenorhabdi
23	158.5	6.6	501	16 Q9KLD3	Q9KLD3 vibrio chol
24	157.5	6.5	485	2 Q53737	Q53737 streptomyce
25	156	6.5	433	17 Q8HRR8	Q8HRR8 halobacteri
26	155.5	6.4	501	2 P61527	P61527 vibrio chol
27	151	6.2	536	16 Q8PFH7	Q8PFH7 xanthomonas
28	150.5	6.2	529	16 Q9PF58	Q9PF58 xylella fas
29	149.5	6.2	606	10 Q94JH3	Q94JH3 oryza sativ
30	149.5	6.2	609	2 Q8VUS4	Q8VUS4 alteromonas
31	149	6.2	500	16 P96264	P96264 mycobacteri
32	147.5	6.1	441	17 Q9HNP7	Q9HNP7 halobacteri
33	146	6.0	430	16 Q9AJU5	Q9AJU5 caulobacter
34	145	6.0	536	16 Q8P326	Q8P326 xanthomonas
35	140	5.8	488	17 Q9UXB4	Q9UXB4 sulfolobus
36	138.5	5.7	374	3 Q04033	Q04033 saccharomyc
37	138.5	5.7	594	16 Q8PFF6	Q8PFF6 xanthomonas
38	137.5	5.7	621	10 Q94JH4	Q94JH4 oryza sativ
39	134	5.5	324	16 Q9F2X2	Q9F2X2 streptomyce
40	134	5.5	768	6 Q8H2V3	Q8H2V3 sus scrofa
41	133.5	5.5	384	3 Q8WZH8	Q8WZH8 agarius bi
42	130	5.4	469	16 Q99ZU1	Q99ZU1 streptococc
43	130	5.4	469	16 Q8K716	Q8K716 streptococc
44	129	5.3	280	10 Q9X428	Q9X428 arabidopsis
45	129	5.3	469	16 Q8P162	Q8P162 streptococc

## ALIGNMENTS

RESULT 1  
ID Q9Y646 PRELIMINARY; PRT; 472 AA.  
AC Q9Y646;

DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Aminopeptidase (Plasma glutamate carboxypeptidase).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Liu C. H., Lin B. Y., Chang L. Y.;  
RT "Cloning of the human aminopeptidase gene."  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RA Strauberg R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF107834; AAD3214.1; -;  
DR EMBL; BC020689; AAH20689.1; -;  
DR MEROPS; M28.014; -;  
DR InterPro; IPR003137; PA.  
DR PROSITE; PS50840; PA.1.  
KW Carboxypeptidase.  
SQ SEQUENCE 472 AA; 51887 MW; EB6CBD2149E042BF CRC64;

Query Match 100.0%; Score 2417; DB 4; Length 472;  
Best Local Similarity 100.0%; Pred. No. 3.6e-176;  
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAFGCVHLLSLCSGKAICKNGISKRTFEIRIEEIASCGDVAKAIINLAVYKAQ 60  
DB 1 MKFLIFAFGCVHLLSLCSGKAICKNGISKRTFEIRIEEIASCGDVAKAIINLAVYKAQ 60

QY 61 NRSYERLALLVDVTGPRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPHMERGESBA 120  
DB 61 NRSYERLALLVDVTGPRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPHMERGESBA 120  
QY 121 VMLERIRIKIALILGSSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVYNOPIYNY 180  
DB 121 VMLERIRIKIALILGSSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVYNOPIYNY 180  
QY 181 SRTVOYRTQGAVEAKVAGALASLIRSASFISYSPHTIGIYQODGVPKIPRACTIVEBAE 240  
DB 181 SRTVOYRTQGAVEAKVAGALASLIRSASFISYSPHTIGIYQODGVPKIPRACTIVEBAE 240  
QY 241 MMSRMAHSGIKIYIOLKMGAKATYPTDTSFNTVAETISGSKPEQVVLVSGHLSMDVQGA 300  
DB 241 MMSRMAHSGIKIYIOLKMGAKATYPTDTSFNTVAETISGSKPEQVVLVSGHLSMDVQGA 300  
QY 301 MDDGGAFISWEALSILKDLGRPKRTLRLVMTAEBOGVGAFQYQOLHKVNISNYSLV 360  
DB 301 MDDGGAFISWEALSILKDLGRPKRTLRLVMTAEBOGVGAFQYQOLHKVNISNYSLV 360  
QY 361 MESDAGTFLPTGLQFTSGSEKARAIIMEEYMSLLOPLNITQVLSHGEGTDINFWIOAGVGA 420  
DB 361 MESDAGTFLPTGLQFTSGSEKARAIIMEEYMSLLOPLNITQVLSHGEGTDINFWIOAGVGA 420  
QY 421 SLDDLKYYKFFPHSHGDTMTVMDPKOMNVAANAVALVSYVADMEMLPRS 472  
DB 421 SLDDLKYYKFFPHSHGDTMTVMDPKOMNVAANAVALVSYVADMEMLPRS 472

## RESULT 2

Q8NBZ1 PRELIMINARY; PRT; 472 AA.  
AC Q8NBZ1. PRELIMINARY; PRT; 472 AA.  
BT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Hypothetical protein FLJ90651.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Tissue=Placenta;  
RA Isegai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,  
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahata K.,  
RA Masuho Y., Ota T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,  
RA Hachori A., Okumura K., Iwayanagi T., Ninomiya K.;  
RT "NEO human cDNA sequencing project."  
RL Submitted (MAR-2002) to the EMBL/genbank/DDbJ databases.  
DR EMBL, AK075132; BAC11423.1; --  
DR PROSITE, IPRO03137; PA.  
DR INTERPRO, PS50840; PA; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 472 AA; 51784 MW; 83F25996F5DC52F3 CRC64;

Query Match 99.4%; Score 2402; DB 4; Length 472;  
Best Local Similarity 99.4%; Pred. No. 5e-175;  
Matches 469; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKFLIFAFPGVHLLSLCSGKAICNGISKRTPEIKERIASCGVAKAIINLAVYGAQ 60  
DB 1 MKFLIFAFPGVHLLSLCSGKAICNGISKRTPEIKERIASCGVAKAIINLAVYGAQ 60  
QY 61 NRSYERLALLVDVTGPRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPHMERGESBA 120  
DB 61 NRSYERLALLVDVTGPRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPHMERGESBA 120  
QY 121 VMLERIRIKIALILGSSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVYNOPIYNY 180  
DB 121 VMLERIRIKIALILGSSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVYNOPIYNY 180  
QY 181 SRTVOYRTQGAVEAKVAGALASLIRSASFISYSPHTIGIYQODGVPKIPRACTIVEBAE 240  
DB 181 SRTVOYRTQGAVEAKVAGALASLIRSASFISYSPHTIGIYQODGVPKIPRACTIVEBAE 240

QY 181 SRTVOYRTQGAVEAKVAGALASLIRSASFISYSPHTIGIYQODGVPKIPRACTIVEBAE 240  
DB 181 SRTVOYRTQGAVEAKVAGALASLIRSASFISYSPHTIGIYQODGVPKIPRACTIVEBAE 240  
QY 241 MMSRMAHSGIKIYIOLKMGAKATYPTDTSFNTVAETISGSKPEQVVLVSGHLSMDVQGA 300  
DB 241 MMSRMAHSGIKIYIOLKMGAKATYPTDTSFNTVAETISGSKPEQVVLVSGHLSMDVQGA 300  
QY 301 MDDGGAFISWEALSILKDLGRPKRTLRLVMTAEBOGVGAFQYQOLHKVNISNYSLV 360  
DB 301 MDDGGAFISWEALSILKDLGRPKRTLRLVMTAEBOGVGAFQYQOLHKVNISNYSLV 360  
QY 361 MESDAGTFLPTGLQFTSGSEKARAIIMEEYMSLLOPLNITQVLSHGEGTDINFWIOAGVGA 420  
DB 361 MESDAGTFLPTGLQFTSGSEKARAIIMEEYMSLLOPLNITQVLSHGEGTDINFWIOAGVGA 420  
QY 421 SLDDLKYYKFFPHSHGDTMTVMDPKOMNVAANAVALVSYVADMEMLPRS 472  
DB 421 SLDDLKYYKFFPHSHGDTMTVMDPKOMNVAANAVALVSYVADMEMLPRS 472

## RESULT 3

Q9Y5X6 PRELIMINARY; PRT; 541 AA.  
AC Q9Y5X6. PRELIMINARY; PRT; 541 AA.  
BT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Blood plasma glutamate carboxypeptidase precursor (EC 3.4.17.21).  
GN PGCP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99223495; PubMed=10206990;  
RA Gingras R., Richard C., El-Alfy M., Morales C.R., Potier M.,  
RA Peshetsky A.V.;  
RT "Purification, cDNA cloning, and expression of a new human blood  
RT plasma glutamate carboxypeptidase homologous to N-acetyl-aspartyl-  
RT alpha-glutamate carboxypeptidase/prolactin-specific membrane antigen."  
RL J. Biol. Chem. 274:11742-11750 (1999).  
DR EMBL, AF119386; AAD31418.1; --  
DR MEROPS, M28.014; --  
DR INTERPRO, IPRO03137; PA.  
DR PROSITE, PS50840; PA; 1.  
KW Carboxypeptidase; Hydrolase; Signal.  
FT SIGNAL 1 24  
FT CHAIN 45 541 BLOOD PLASMA GLUTAMATE CARBOXYPEPTIDASE.  
SQ SEQUENCE 541 AA; 59931 MW; 9B138FES7081F68 CRC64;

Query Match 94.5%; Score 2285; DB 4; Length 541;  
Best Local Similarity 99.6%; Pred. No. 5.4e-166;  
Matches 445; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKFLIFAFPGVHLLSLCSGKAICNGISKRTPEIKERIASCGVAKAIINLAVYGAQ 60  
DB 1 MKFLIFAFPGVHLLSLCSGKAICNGISKRTPEIKERIASCGVAKAIINLAVYGAQ 60  
QY 61 NRSYERLALLVDVTGPRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPHMERGESBA 120  
DB 61 NRSYERLALLVDVTGPRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPHMERGESBA 120  
QY 121 VMLERIRIKIALILGSSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVYNOPIYNY 180  
DB 121 VMLERIRIKIALILGSSSIGTPPEGITAEVLVVTSPFDELQRRASBARSKIIVYNOPIYNY 180  
QY 181 SRTVOYRTQGAVEAKVAGALASLIRSASFISYSPHTIGIYQODGVPKIPRACTIVEBAE 240  
DB 181 SRTVOYRTQGAVEAKVAGALASLIRSASFISYSPHTIGIYQODGVPKIPRACTIVEBAE 240  
QY 241 MMSRMAHSGIKIYIOLKMGAKATYPTDTSFNTVAETISGSKPEQVVLVSGHLSMDVQGA 300  
DB 241 MMSRMAHSGIKIYIOLKMGAKATYPTDTSFNTVAETISGSKPEQVVLVSGHLSMDVQGA 300

```

Db 241 MMSHMGHGIYIVQLKMGAKTYPDTSFNTVAETGTSKYPEQVVLVSGHDSMDVCGA 300
Qy 301 MDDGGAFISWEALSLIKDGLRPRKTRLVMTAEEGGVGAFOYVOLHVNINSNLSV 360
Db 301 MDDGGAFISWEALSLIKDGLRPRKTRLVMTAEEGGVGAFOYVOLHVNINSNLSV 360
Qy 361 MESDAGTFLPTGLQFTGSEKARAIMEEVMSLLOPLNTQVLSHGEGTDINFIQAGVGA 420
Db 361 MESDAGTFLPTGLQFTGSEKARAIMEEVMSLLOPLNTQVLSHGEGTDINFIQAGVGA 420
Qy 421 SLDDLYKYFFHHSHGDTMTVMDPKQ 447
Db 421 SLDDLYKYFFHHSHGDTMTVMDPSR 447

```

## RESULT 4

```

Q9WVJ3 PRELIMINARY; PRT; 470 AA.
ID Q9WVJ3;
AC Q9WVJ3;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Amino peptidase (Similar to plasma glutamate carboxypeptidase).
GN PGCP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Liu C.H., Lin B.Y., Chang L.Y.;
RT "Cloning of the mouse aminopeptidase gene.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strassberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body, and Mesonephros;
RX MEDLINE=22354683; PubMed=12466851;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AF107835; AA043215.1; -
DR EMBL; BC037067; AA037067.1; -
DR EMBL; AK032972; BAC28105.1; -
DR EMBL; AK075666; BAC35891.1; -
DR MEROPS; M28.014; -
DR MGD; MGI:1889205; PgcP.
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR001337; PA.
DR PROSITE; PS00223; ANNEXIN; 1.
DR PROSITE; PS00840; PA; 1.
KM Carboxypeptidase.
SQ SEQUENCE 470 AA; 51813 MW; 0F3490681691866A CRC64;

```

Query Match 88.2%; Score 2133; DB 11; Length 470;  
 Best Local Similarity 87.7%; Pred. No. 1,8e-154;  
 Matches 414; Conservative 28; Mismatches 28; Indels 2; Gaps 1;

```

Qy 1 MKFLIFAFGGVHLSTCSGKAICNGISKRTFEIEKEIASCGDVAKAIINLAVYGAQ 60
Db 1 MRSFLFFLVF--VHLALSGKAVFRNGVSGRTFEIEKEIANEYDAKAIINLAVYGRQ 58
Qy 61 NRSYERLALVDVTPGPRISGSKNLEKAIQIMYQNLQODGLEKHLPEVRIIPMERGES 120
Db 59 NRSYERLGLVDVTPGPRISGSKNLEKAIQIMYQNLQODGLEKHLPEVRIIPMERGES 118

```

```

Qy 121 VMLPRILHKIALLIGLSSIGTPPEGITAEVLVWTSPELORRASEARGKIIVYNQPYNY 180
Db 116 VMLPRILHKIALLIGLSSIGTPPEGITAEVLVWTSPELORRASEARGKIIVYNQPTGX 178
Qy 181 SRTVQYRTQGAVEAKVAGALASLRSVASFISPHGTIOEYODGVKPIPTACTIVDAE 240
Db 179 EKTQYVAVQGAVEAKVAGALASLRSVASFISPHGTIOEYODGVKPIPTACTIVDAE 238
Qy 241 MMSHMGHGIYIVQLKMGAKTYPDTSFNTVAETGTSKYPEQVVLVSGHDSMDVCGA 300
Db 239 MMSHMGHGIYIVQLKMGAKTYPDTSFNTVAETGTSKYPEQVVLVSGHDSMDVCGA 298
Qy 301 MDDGGAFISWEALSLIKDGLRPRKTRLVMTAEEGGVGAFOYVOLHVNINSNLSV 360
Db 299 LDDGGAFISWEALSLIKDGLRPRKTRLVMTAEEGGVGAFOYVOLHVNINSNLSV 358
Qy 361 MESDAGTFLPTGLQFTGSEKARAIMEEVMSLLOPLNTQVLSHGEGTDINFIQAGVGA 420
Db 359 MESDAGTFLPTGLQFTGSEKARAIMEEVMSLLOPLNTQVLSHGEGTDINFIQAGVGA 418
Qy 421 SLDDLYKYFFHHSHGDTMTVMDPKQNTAAAVAVSYVADMEMLPRS 472
Db 419 SLDDLYKYFFHHSHGDTMTVMDPKQNTAAAVAVSYVADMEMLPRS 470

```

## RESULT 5

```

Q9JLV0 PRELIMINARY; PRT; 472 AA.
ID Q9JLV0;
AC Q9JLV0;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Liver annexin-like protein.
GN LAL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Servillo G., Della Fazio M.A., Piobbico D., Bartoli D., Castellani M.,
RA Brancorsini S., Viola Magni M.;
RT "LAL, a novel gene involved during liver regeneration.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131077; AAF36518.1; -
DR MEROPS; M28.014; -
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR001337; PA.
DR PROSITE; PS00223; ANNEXIN; 1.
DR PROSITE; PS00840; PA; 1.
SQ SEQUENCE 472 AA; 52011 MW; 55C68FB90E63F265 CRC64;

```

Query Match 87.6%; Score 2117; DB 11; Length 472;  
 Best Local Similarity 87.5%; Pred. No. 3e-153;  
 Matches 413; Conservative 24; Mismatches 35; Indels 0; Gaps 0;

```

Qy 1 MKFLIFAFGGVHLSTCSGKAICNGISKRTFEIEKEIASCGDVAKAIINLAVYGAQ 60
Db 1 MRFLLFFLVAVVHLFSLGSGKAIFYSGVSGRTFOEIEKEIANEYDAKAIINLAVYGRQ 60
Qy 61 NRSYERLALVDVTPGPRISGSKNLEKAIQIMYQNLQODGLEKHLPEVRIIPMERGES 120
Db 61 NRSYERLGLVDVTPGPRISGSKNLEKAIQIMYQNLQODGLEKHLPEVRIIPMERGES 120
Qy 121 VMLPRILHKIALLIGLSSIGTPPEGITAEVLVWTSPELORRASEARGKIIVYNQPYNY 180
Db 121 VMLPRILHKIALLIGLSSIGTPPEGITAEVLVWTSPELORRASEARGKIIVYNQPTDX 180
Qy 181 SRTVQYRTQGAVEAKVAGALASLRSVASFISPHGTIOEYODGVKPIPTACTIVDAE 240
Db 181 GKTQYVERGAVEAKVAGALASLRSVASFISPHGTIOEYODGVKPIPTACTITIDAE 240

```

```
Qy 241 MMSHMASHGIXIVIQIKMGAKTYPTDSEFNTVAETGSKYPEQVVLVSGHLSMDVGGGA 300
Db 241 MMSHMASHGIXIVIHLMKGAATYPTDSEFNTVAETGSKYPEQVVLVSGHLSMDVGGGA 300
Qy 301 MDGGGAFISWEALSLIKDLGRLPKRTLRVLTAEEOGGVGAFOYYOLHKVNISNYSVLV 360
Db 301 LDDGGGAFISWEALSLIKDLGRLPKRTLRVLTAEEOGGVGAFOYYOLHKVNISNYSVLV 360
Qy 361 MESDAGTFLPTGLQFTSGSEKARAIMBEVMSLQPLNITQVLSHGEGTDINFWIOAGVPGA 420
Db 361 MEADSGTFLPTGLQFTSGSEKARAIMBEVMSLQPLNITQVLSHGEGTDINFWIOAGVPGA 420
Qy 421 SLDDLYKYFFHHSHGDTMTVMDPKQNVAAVAAVAVSVVADMEMLPRS 472
Db 421 SLRDDLYKYFFHHSHGDTMTAMDPKQNVAAVAAVAVAVVADMEMLPRS 472
```

## RESULT 6

```
Q921Y1 PRELIMINARY; PRT; 472 AA.
ID 0921Y1
AC 0921Y1
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hematopoietic lineage switch 2 related protein.
GN Hls2-RP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCB1_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Flisher;
RA Chen Y., Talmage D.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF097723; AAC72384.1; -.
DR MEROPS; M28.014; -.
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR001337; PA.
DR PROSITE; PS00223; ANNEXIN; 1.
DR PROSITE; PS50840; PA; 1.
SQ SEQUENCE 472 AA; 51969 MW; DC211165187956E2 CRC64;
```

Query Match 87.5%; Score 2116; DB 11; Length 472;  
Best Local Similarity 87.5%; Pred. No. 3.6e-153;  
Matches 413; Conservative 24; Mismatches 35; Indels 0; Gaps 0;

```
Qy 1 MKFLIFAFGGVHLISLCSGKAIKNGISKRTFEEIKEEIASCGDVAKAIINLAVYGRQAQ 60
Db 1 MRFLFELVAVVHLFSLSGKAIYKSGVSGRTFQKEIANYEDVAKAIINLAVYGRKYQ 60
Qy 61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHLEPRIPMERGESSA 120
Db 61 NRSYERLGLVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHLEPRIPMERGESSA 120
Qy 121 VMLBPRIHKAIIIGLGGSSIGTPPEGITAEVLVNTSPDELQRRASARGKIVVYNQPYINY 180
Db 121 VMLBPRIHKAIIIGLGGSSIGTPPEGITAEVLVNTSPDELQRRASARGKIVVYNQPYINY 180
Qy 181 SRTVOYRTQGAWEAKVAGALASLIRSVASFISYPHTGIOEYODGVPIPTACTIVEDAE 240
Db 181 GKTVOYRGAWEAKVAGALASLIRSVASFISYPHTGIOEYODGVPIPTACTIVEDAE 240
Qy 241 MMSHMASHGIXIVIQIKMGAKTYPTDSEFNTVAETGSKYPEQVVLVSGHLSMDVGGGA 300
Db 241 MMSHMASHGIXIVIHLMKGAATYPTDSEFNTVAETGSKYPEQVVLVSGHLSMDVGGGA 300
Qy 301 MDGGGAFISWEALSLIKDLGRLPKRTLRVLTAEEOGGVGAFOYYOLHKVNISNYSVLV 360
Db 301 LDDGGGAFISWEALSLIKDLGRLPKRTLRVLTAEEOGGVGAFOYYOLHKVNISNYSVLV 360
Qy 361 MESDAGTFLPTGLQFTSGSEKARAIMBEVMSLQPLNITQVLSHGEGTDINFWIOAGVPGA 420
Db 361 MEADSGTFLPTGLQFTSGSEKARAIMBEVMSLQPLNITQVLSHGEGTDINFWIOAGVPGA 420
```

```
Db 361 MEADSGTFLPTGLQFTSGSEKARAIMBEVMSLQPLNITQVLSHGEGTDINFWIOAGVPGA 420
Qy 421 SLDDLYKYFFHHSHGDTMTVMDPKQNVAAVAAVAVSVVADMEMLPRS 472
Db 421 SLRDDLYKYFFHHSHGDTMTAMDPKQNVAAVAAVAVAVVADMEMLPRS 472
```

## RESULT 7

```
Q70216 PRELIMINARY; PRT; 433 AA.
ID 070216
AC 070216
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hematopoietic lineage switch 2.
GN PCGP OR Hls2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Williams J.H., Chan C.-Y., Klinken S.P.;
RT "Hematopoietic lineage switch 2 (Hls2), a novel mRNA species induced during an erythroid to myeloid lineage switch."
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF009513; AAC17945.1; -.
DR MEROPS; M28.014; -.
DR MGD; MGI:1889205; PgcP.
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR001337; PA.
DR PROSITE; PS00223; ANNEXIN; 1.
DR PROSITE; PS50840; PA; 1.
SQ SEQUENCE 433 AA; 47858 MW; 89C9B93712004669 CRC64;
```

Query Match 78.9%; Score 1907; DB 11; Length 433;  
Best Local Similarity 84.8%; Pred. No. 2.9e-137;  
Matches 374; Conservative 25; Mismatches 28; Indels 14; Gaps 2;

```
Qy 1 MKFLIFAFGGVHLISLCSGKAIKNGISKRTFEEIKEEIASCGDVAKAIINLAVYGRQAQ 60
Db 1 MRFLFELVAVVHLFSLSGKAIYKSGVSGRTFQKEIANYEDVAKAIINLAVYGRKYQ 60
Qy 61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHLEPRIPMERGESSA 120
Db 61 NRSYERLGLVDYVGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHLEPRIPMERGESSA 118
Qy 181 SRTVOYRTQGAWEAKVAGALASLIRSVASFISYPHTGIOEYODGVPIPTACTIVEDAE 240
Db 179 EKTVOYRGAWEAKVAGALASLIRSVASFISYPHTGIOEYODGVPIPTACTIVEDAE 238
Qy 241 MMSHMASHGIXIVIQIKMGAKTYPTDSEFNTVAETGSKYPEQVVLVSGHLSMDVGGGA 300
Db 239 MMSHMASRGKIVIHLMKGAATYPTDSEFNTVAETGSMPE-----DVGGGA 286
Qy 301 MDGGGAFISWEALSLIKDLGRLPKRTLRVLTAEEOGGVGAFOYYOLHKVNISNYSVLV 360
Db 287 LDDGGGAFISWEALSLIKDLGRLPKRTLRVLTAEEOGGVGAFOYYOLHKVNISNYSVLV 346
Qy 361 MESDAGTFLPTGLQFTSGSEKARAIMBEVMSLQPLNITQVLSHGEGTDINFWIOAGVPGA 420
Db 347 MEADSGTFLPTGLQFTSGSEKARAIMBEVMSLQPLNITQVLSHGEGTDINFWIOAGVPGA 406
Qy 421 SLDDLYKYFFHHSHGDTMT 441
Db 407 SLRDDLYKYFFHHSHGDTMT 427
```

## RESULT 8

076552 ID 076552 PRELIMINARY; PRT; 493 AA.  
 AC 076552;  
 DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Aminopeptidase ES-62 precursor.  
 OS Acanthocheilomena viteae (Dipetalonema viteae).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Splirurida; Filarioidae;  
 OC Onchocercidae; Acanthocheilomena.  
 NCBI\_TaxID=6277;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20055923; PubMed=10589978;  
 RA Harnett W., Houston K.M., Tate E., Garate T., Apfel H., Adam R.,  
 RA Hellem S.M., Panico M., Paxton T., Dell A., Morris H., Brzecki H.,  
 RT "Molecular cloning and demonstration of an aminopeptidase activity in  
 RT a filarial nematode glycoprotein."  
 RL Mol. Biochem. Parasitol. 104:11-23(1999).  
 DR EMBL; AF077194; AAC28365.1; -.  
 DR MEROPS; M28.015; -.  
 DR InterPro; IPR003137; PA.  
 DR PROSITE; PS50840; PA; 1.  
 KW Signal.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 493 AMINOPEPTIDASE ES-62.  
 SQ SEQUENCE 493 AA; 54916 MW; C7E9227DD18B1DF4 CRC64;  
 Query Match 31.4%; Score 759.5; DB 5; Length 493;  
 Best Local Similarity 36.6%; Pred. No. 1.7e-49;  
 Matches 174; Conservative 92; Mismatches 166; Indels 43; Gaps 13;  
 QY 15 LSLGSGKAIC-----KNGSKRTFEIEKEIASCGVAKAIIINLVYGAQRYSYERL 67  
 DB 14 LTVVLGAIVLPDKTVPAPKNYIQETPGKEVAE-----LQYITKGEVGLAYOWL 62  
 QY 68 ALIVDTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLPEY-RIPMERGESAVMLEPR 126  
 DB 63 SKLVDSGCHNVGSDLSLEKTAIFLESKXNDNPFKHEEVPNLPHWRKNDVEMIEPR 122  
 QY 127 IHKTAIIGLSSISITPREGITAEVLVYTSFDELORRASEARGKIVVNOPIYNSRTVQY 186  
 DB 123 NQRLNVAIG---GSEPASATGEVTVIYDLDDV--KPDVDVAGKIVTQTFAGYPLTLKY 177  
 QY 187 RTQGAVEAAKYGALASLIRSVASFSIYSPHTGIEQYQGVKIPACTIVDEAMSGMA 246  
 DB 178 R-RSVKLVFEQJGALGVLVKISITFSINSPTGTGAEN--TTTPACLTIEAEMLERLY 233  
 QY 247 SHGIKIVIQLMKGAATYPTDTSFNTVAEITGSKYPEQVVLVSGHLSMDVQAGAMDGGG 306  
 DB 234 RSGKKIVRMKMKHYEERPINSNLIFRITGSERPSEVVLISAVDSMDVQAGALDDGAG 293  
 QY 307 AFISWEALSLIKDGL---LRPKRTLRVLMTAEBOGVGAFOYYQLHKVNI-SNYSLV 360  
 DB 294 CAVVMSALHSIKLAERPKFKPKRTIRIGIFMTSEBOGVGAKHYIYHKNDSEKFEYFV 353  
 QY 361 MESDAGTLPFG---LQFTGSEKARAMEEVMSLQDLANTQVL--SHSGGTINFPIQ 414  
 DB 354 SETPTGFKSTNMLAHSFSGDKKSMLEKIRLISNGALGINSVVG--DVTFFAK 412  
 QY 415 AGVFGASLDD--LYKYFFPHSHSDTVMDDPKOMVAAVMAVSVVADMEB 467  
 DB 413 DGISVNTIPKAVDYVYFFHHTAGDVMTVLKDQDLETTTISFATLGAVIANMD 467

## RESULT 9

QY UNMB PRELIMINARY; PRT; 140 AA.  
 AC QYUNMB;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Aminopeptidase (Fragment).

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Liu C.H., Lin B.Y., Chang L.Y.;  
 RT "Cloning of the human aminopeptidase gene."  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF107833; AAD43213.1; -.  
 DR MEROPS; M28.014; -.  
 DR InterPro; IPR003137; PA.  
 DR PROSITE; PS50840; PA; 1.  
 FT NON\_TER 1 1  
 FT SEQUENCE 140 AA; 15472 MW; F34ACEAB33A1AD24 CRC64;  
 Query Match 29.0%; Score 701; DB 4; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-46;  
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LSGSKNLEKAIQIMYQNLQODGLEKVLPEYRIPMERGESAVMLEPRIHKTAIILGSG 137  
 DB 1 LSGSKNLEKAIQIMYQNLQODGLEKVLPEYRIPMERGESAVMLEPRIHKTAIILGSG 60  
 QY 138 SIGPREGITAEVLVYTSFDELORRASEARGKIVVNOPIYNSRTVQYTGAVEAKY 197  
 DB 61 SIGPREGITAEVLVYTSFDELORRASEARGKIVVNOPIYNSRTVQYTGAVEAKY 120  
 QY 198 GALASLIRSVASFSIYSPHT 217  
 DB 121 GALASLIRSVASFSIYSPHT 140

## RESULT 10

QY UNMB PRELIMINARY; PRT; 414 AA.  
 AC QYUNMB;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Peptidase, M28D family.  
 GN SO3539.  
 OS Shewanella oneidensis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
 OC Alteromonadaceae; Shewanella.  
 NCBI\_TaxID=70863;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MR-1;  
 RX MEDLINE=22297686; PubMed=12368813;  
 RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,  
 RA Read T.D., Eisen J.A., Seehardt R., Ward N., Methe B., Clayton R.A.,  
 RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkak L., Daugherty S.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,  
 RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,  
 RA Venter A., Weidman J., Impraim M., Lee K., Berry K., Lee C.,  
 RA Mueller T.V., Smith H.O., Venter J.C., Neale K.H., Fraser C.M.,  
 RT "Genome sequence of the dissimilatory metal ion-reducing bacterium  
 RT Shewanella oneidensis."  
 RL Nat. Biotechnol. 20:1118-1123(2002).  
 DR EMBL; AE015790; AAN56530.1; -.  
 DR TIGR; SO3539; -.  
 KW Complete proteome.  
 SQ SEQUENCE 414 AA; 44028 MW; C9AEB2A5773DBCB4 CRC64;  
 Query Match 27.1%; Score 655; DB 16; Length 414;  
 Best Local Similarity 36.4%; Pred. No. 1.3e-41;  
 Matches 148; Conservative 72; Mismatches 173; Indels 14; Gaps 4;

QY 74 VGPLSGSKNLEKAIQIMYQNLQODGLEKVLPEYRIPMERGESAVMLEPRIHKTAIIL 133

```

Db      3 VGPRLASPKDVIAYNMAMNKLTLGFDKVKYEPQVPIWMEGSAKAKIISPEQPLVIT 62
Qy      134 GLGSSIGTPPGITAEVLVTSFDELQRA--SEANGKIVTNQPI-----NYRRTQY 186
Db      63 ALGGSVAAPVPGIKKAKIFNSLEALQOATPDVDVGAIAFDQKTVRHITGEGKSVGG 122
Qy      187 RTQGAWEAKVAGALASLIRSVASFISYSPHTGIEOYQGVPKIPACTIVEDAEWMSMA 246
Db      123 RSKGIAIAAOKGAVAVIIRISIGTDHDMAHGIMYQDGVFKIPAAANSNPADLVADM 182
Qy      247 SHGKIIVQLKMGAKTYPDTDSFNTVAETISKYPEOVVLVSGHLDSDVDGAMDDGG 306
Db      183 KRDPNAVLELNNSPKDLGNTSYNVIAEYTGSSKNEIYLIGAHLDSDDEGTGALDDG 242
Qy      307 AFISWEALSLIKDLRPRKTRILVMTAREOGVGAFQYQYLHKVNISNYSLVMSDAG 366
Db      243 IAIYAAAAKHIDLPQKAPRTIRVLYAAEEMGLIGKAYAEAHKTELPLHYIAESFPG 302
Qy      367 TFLPTGLQFTGSEKARAIMEEVMSLLOPLNITQVL-----SHGEGTDINFWIQAGVPGASL 422
Db      303 AGPIYQITTKVNEK--VFAQVQESIKPMYNGVALGDNQASGDPISMLPALGVVASL 359
Qy      423 LDDLKYEFFHSHGDTMTVMDPKQNNVAAAAMVAVSVYVADMEEML 469
Db      360 RODGRDYDFYHHTPNDTLDKINPKALQAVAAQFAVYIMANSIVL 406

```

## RESULT 11

```

Q9A5A8      PRELIMINARY;      PRT;      467 AA.
ID  Q9A5A8;
AC  01-JUN-2001 (TREMBLrel. 17, Created)
DT  01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT  01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE  Aminoacidase, putative.
GN  CC2544.
OS  Caulobacter crescentus.
OC  Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC  Caulobacteraceae; Caulobacter.
OX  NCBI_TaxID=155892;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=ATCC 19089 / CB15;
RX  MEDLINE=21173698; PubMed=11259647;
RA  Nieman W.C., Fejdllyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA  Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA  Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA  DeBoy R.T., Dodson R.J., Durkin A.S., Gwynn M.L., Haft D.H.,
RA  Kolonay J.F., Smit J., Craven M.B., Khouri H., Shertly J., Berry K.,
RA  Ueberbach T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA  Salzberg S.L., Venter J.C., Shapiro L., Frazer C.M.;
RT  "Complete genome sequence of Caulobacter crescentus.";
RL  Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).
DR  EMBL; AE005922; AAK24515.1; -.
DR  TIGR; CC2544; -.
DR  InterPro; IPR003137; PA.
DR  PROSITE; PS00840; PA. 1.
KW  Complete proteome.
SQ  SEQUENCE 467 AA; 48918 MW; 9984DE2A9D84F00F CRC64;

```

Query Match 26.5%; Score 640; DB 16; Length 467;

Best Local Similarity 36.2%, Pred. No. 2.2e-40; Indels 10; Gaps 4;

Matches 147; Conservative 72; Mismatches 177;

```

Qy      70 LVDYGPRLSGSKNLEKAIQIMYQNLQOGEKLEHLEPRIPIMWEGESAVMLEPRITK 129
Db      49 LTTNIGPRLVSGPAAKAKDMSVAKFALGFNINIKVDEAKKSWGEGSSALVAPYMK 108
Qy      130 IAILGLSSIGTPPGITAEVLVTSFDELQRASEA-RGKIVVYNQPIYNSRTVQY-- 186
Db      109 LGAVGLGRVSTPAGIGIEAEVALFTFTADWMAAPDGAALGKIVITQPVWRTONGAGYGA 168

```

```

Qy      187 ----RTQGAWEAKVAGALASLIRSVASFISYSPHTGIEOYQGVPKIPACTIVEDAEWM 242
Db      169 AGVSRRIRGPVBAARAGVAMILIRISITSDSVPHPTGTGASEGVVTPAALIGVBAEQ 228
Qy      243 SRMASHGKIIVQLKMGAKTYPDTDSFNTVAETISKYPEOVVLVSGHLDSDVDGQAM 302
Db      229 ERLARKVPMKVKKLSSVPPNNVANNISGIDIGSEKPEVEIVYIGHLDSWDVGTGALD 288
Qy      303 DGGAFISWEALSLIKDLRPRKTRILVMTAREOGVGAFQYQYLHKVNISNYSLVME 362
Db      289 DATGIALTAALAKLIGDLPKRPKRTIRVMMGSESG--SEAYLAANKALSTMVLAGE 347
Qy      363 SDAGTFLPTGLQFTGSEKARAIMEEVMSLLOPLNITQVL-----TOVLSHGETDINFWIQAGVGA 420
Db      348 SDTGADRIYSLQVPAGSLDHHVVAASVLAFLKTIYDRTPAAGADVSGIERAGVPVI 407
Qy      421 SLDDLKYEFFHSHGDTMTVMDPKQNNVAAAAMVAVSVYVADM 466
Db      408 NINQDASRYPDYHHTMDTLDKVPAPALQAVAAWTSILVLYLVADSD 453

```

## RESULT 12

```

Q9R114      PRELIMINARY;      PRT;      139 AA.
ID  Q9R114;
AC  01-MAY-2000 (TREMBLrel. 13, Created)
DT  01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT  01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE  Aminoacidase (fragment).
GN  PGCP.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Liver;
RA  Liu C.H., Lin B.Y., Chang L.Y.;
RT  "Cloning of the mouse aminoacidase gene.";
RL  Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF107832; AAD43212.1; -.
DR  HSSP; P80561; 1X00.
DR  MEROPS; M28.014; -.
DR  MGD; MGI:1889205; Pgcp.
FT  NON_TER 1
FT  NON_TER 139
SQ  SEQUENCE 139 AA; 15161 MW; 3B21B7CF2D2DE648 CRC64;

```

Query Match 25.8%; Score 623; DB 11; Length 139;

Best Local Similarity 84.7%, Pred. No. 6.2e-40; Indels 0; Gaps 0;

Matches 116; Conservative 14; Mismatches 7; Indels 0;

```

Qy      277 GSKTPEOVVLVSGHLDSDVDGQAMDDGGAFFISWEALSLIKDLRPRKTRILVLTAE 336
Db      3 GSMYPEEVVLVSGHLDSDVDGQALDDGGGFISWEALSLIKDLRPRKTRILVLTAE 62
Qy      337 EOGGVGAFOYQYLHKVNISNYSLVMSDAGTFLPTGLQFTGSEKARAIMEEVMSLLOPLN 396
Db      63 EOGGIGASQYELHKANISKSLVMEADSGTFLPTGLQFTGSDKARAIMEEVMSLLOPLN 122
Qy      397 ITQVLSHGETDINFWI 413
Db      123 VTKVFSNCGEIDISFRI 139

```

## RESULT 13

```

Q9PHE8      PRELIMINARY;      PRT;      472 AA.
ID  Q9PHE8;
AC  08PHE8;
DT  01-OCT-2002 (TREMBLrel. 22, Created)
DT  01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT  01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE  Aminoacidase.

```

GN XAC3309.  
 OS Xanthomonas axonopodis (pv. citri).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 CX NCBI\_TaxID=92829;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=306 / ATCC 13902 / XV 101;  
 RX MEDLINE=22022145; PubMed=12024217;  
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
 RA Camarotte G., Camnavan F., Cardozo J., Chamberggo F., Clapina L.P.,  
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truifi D., Tsai S.M., White F.F.,  
 RA Setubal J.C., Kitajima J.P.;  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities.";  
 RL Nature 417:459-463(2002).  
 DR EMBL: AEO11976; AAA38152.1; -  
 DR InterPro: IPR001337; PA.  
 DR PROSITE: PS50840; PA; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 472 AA; 48788 MW; 9956CDF1F49FC68 CRC64;  
  
 Query Match 20.7%; Score 499.5; DB 16; Length 472;  
 Best Local Similarity 29.7%; Pred. No. 1.2e-29;  
 Matches 140; Conservative 76; Mismatches 218; Indels 37; Gaps 10;  
  
 QY 12 VHLTLSCGKAIKNGI---SKRTFEIEKIEIACGVAKAIIIMAVYKQNSYERLA 68  
 Db 8 ISALLASCALAAATSIIPDSALRTAQLREQ-ALADDTGRAVVO----- 50  
 QY 69 LLDVTVGPRLSGSKNLEKAIQIMYQNLQDGLKRVHLEPVRIIPMERGESAVMLEPRRIH 128  
 Db 51 SLTFTEVGPRIAGBADPRAVAMAKAFASLGFCDKVTETPTFPKMERSEHAIVGAHAQ 110  
 QY 129 KIALIGSSIGTPPEGITAEVLVVTSPDELQRR-ASEARKIVVYNQPYI-----NYS 181  
 Db 111 PLTITLGGSPGGTVEG---EVRFEETLAAQLAAPASGLAKIAFVYQWTKARDGXDYG 167  
 QY 182 RTVOYRTQGAVEAKVGAALSLRSVASFISYSPHTGIGQYQDGVPKITPACITVEDAEM 241  
 Db 168 NGCAVRSKSGSEAIRKGAIGFVMSAGTDSHRVPHGTITTFDEBLTFVPAALSLVPDNO 227  
 QY 242 MSRWASHGIKIVIOIKMGAKTYPDTDSFNTVAEITGSKYEPQVVLVSGHLSMDVGOGAM 301  
 Db 228 LARITLALG-SARVRLALDCCMGDTATSYNVIGETIGRSKREVVVIGHLSMDLGTGAI 286  
 QY 302 DDGGAATISWEALSLIKDLGIRPKRTLRVLTAEBOGCVGAFQYQLH---KYNISNYS 358  
 Db 287 DDGGAATITMAAGHLIQQLKQAPKRTIRVAVAFANEBOGLHGKAYAAHGDADMDLHQ 346  
 QY 359 LVMSDD--AGTFPLTGLOFTGSEKARAIMEEVMSLLOPLINTQYLSH-GESTDIPNFIQA 415  
 Db 347 IGASDVGAGRIYAFNTGAAPDSDRAATKQIAEVLAPLGIAYEPSSKGGPDPVGPISAK 406  
 QY 416 GVPASLILDLKYYFFPHSHSGDTMTVMDPKQNNVAAAWAVSVVADME 466  
 Db 407 GGAGWGLAQDGTGYFDLHHTADDTLTKIDPKALQNNVAAAYVAVLAAD 457

AC Q8P625;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Xanthomonadaceae.  
 GN XCC3157.  
 OS Xanthomonas campestris (pv. campestris).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 CX NCBI\_TaxID=340;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33913 / NCPPB 528;  
 RX MEDLINE=22022145; PubMed=12024217;  
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
 RA Camarotte G., Camnavan F., Cardozo J., Chamberggo F., Clapina L.P.,  
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
 RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truifi D., Tsai S.M., White F.F.,  
 RA Setubal J.C., Kitajima J.P.;  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities.";  
 RL Nature 417:459-463(2002).  
 DR EMBL: AEO12431; AAA42427.1; -  
 DR InterPro: IPR001337; PA.  
 DR PROSITE: PS50840; PA; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 375 AA; 38756 MW; C75F4A4914E2EA05 CRC64;  
  
 Query Match 17.4%; Score 420; DB 16; Length 375;  
 Best Local Similarity 33.0%; Pred. No. 9.9e-24;  
 Matches 105; Conservative 54; Mismatches 145; Indels 14; Gaps 5;  
  
 QY 59 AONRSYERLALVDVTPRLSGSKNLEKAIQIMYQNLQDGLKRVHLEPVRIIPMERGEE 118  
 Db 41 ADSTGFVAVSILTFTEVGPRIAGBADPRAVAMAKAFASLGFCDKVTETPTFPKMERSE 100  
 QY 119 SAVMLEPRIRKIALILGSSIGTPPEGITAEVLVVTSPDELQRR-ASEARKIVVYNQPY 177  
 Db 101 OAAVIGHAQPLHITLALGSPGGTVEG---EIVREFEMLAAQLAAPASGLAKIAFVYQW 157  
 QY 178 I-----NYSRTVOYRTQGAVEAKVGAALSLRSVASFISYSPHTGIGQYQDGVPKIPT 231  
 Db 158 VKARDGDYNGCAVRSKSGSEAIRKGAIGFVMSAGTDSHRVPHGTITTFDEBLTFVPA 217  
 QY 232 ACTIVEDAEEMSWASHGIKIVIOIKMGAKTYPDTDSFNTVAEITGSKYEPQVVLVSGHLS 291  
 Db 218 AALSVPANQALRIALALG-STVRRLALDCCMGDTATSYNVIGETIGRSKREVVVIGHLS 276  
 QY 292 DSDMDVGAGADDGGAATISWEALSLIKDLGIRPKRTLRVLTAEBOGCVGAFQYQLH 351  
 Db 277 DSDMDLGTGALIDDGGAATITMAAGHLIQQLKQAPKRTIRVAVAFANEBOGLHGKAYAAH 336  
 QY 352 VNISNYSLVN---ESDNG 366  
 Db 337 KDAKDMALHQIGASDFG 354

RESULT 14  
 Q8P625  
 ID Q8P625  
 PRELIMINARY; PRT; 375 AA.

RESULT 15  
 Q93EJ5  
 ID Q93EJ5  
 AC Q93EJ5  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:49:10 ; Search time 37.9351 Seconds  
(without alignments)  
526.445 Million cell updates/sec

Title: US-09-745-763-36

Perfect score: 2417  
Sequence: 1 MKFLIFAFGCVHLLSICSG.....AVNAVSVYVADMEMLPRS 472

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PTCUTS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2417	100.0	473	US-09-482-273-120	Sequence 120, App
2	229.5	9.5	496	US-09-079-955-2	Sequence 2, Appli
3	226.5	9.4	481	US-09-330-095-1	Sequence 1, Appli
4	182	7.5	584	US-09-252-991A-18292	Sequence 18292, A
5	165	6.8	33	US-09-482-273-224	Sequence 224, App
6	165	6.8	34	US-09-482-273-187	Sequence 187, App
7	165	6.8	750	US-08-325-553-2	Sequence 2, Appli
8	165	6.8	750	US-08-394-152A-2	Sequence 2, Appli
9	165	6.8	750	US-09-044-668-2	Sequence 2, Appli
10	165	6.8	750	US-08-705-477E-2	Sequence 2, Appli
11	161	6.7	693	US-08-705-477E-101	Sequence 101, App
12	160	6.6	707	US-08-528-122-18	Sequence 18, Appli
13	160	6.6	707	PCT-US95-11720-18	Sequence 18, Appli
14	143.5	5.9	622	US-08-547-197-1	Sequence 1, Appli
15	143.5	5.9	622	US-08-957-940-1	Sequence 1, Appli
16	137.5	5.7	760	US-09-079-955-5	Sequence 5, Appli
17	126	5.2	760	US-08-547-197-2	Sequence 2, Appli
18	126	5.2	760	US-08-957-940-2	Sequence 2, Appli
19	108.5	4.5	439	US-09-634-218-412	Sequence 412, App
20	102.5	4.2	337	US-09-107-532A-5998	Sequence 5998, App
21	100.5	4.2	647	US-09-423-439-60	Sequence 60, Appli
22	100.5	4.2	4928	US-09-036-987A-5	Sequence 5, Appli
23	100.5	4.2	4928	US-09-370-760-5	Sequence 5, Appli
24	100.5	4.2	4928	US-09-603-207-5	Sequence 5, Appli
25	100	4.1	2325	US-08-417-089-6	Sequence 6, Appli
26	100	4.1	2325	US-08-695-651-6	Sequence 6, Appli
27	100	4.1	2325	US-08-930-285-6	Sequence 6, Appli

28	100	4.1	2325	US-08-695-421-6	Sequence 6, Appli
29	100	4.1	2325	US-08-697-826A-10	Sequence 10, Appli
30	98.5	4.1	934	US-09-252-991A-26779	Sequence 26779, A
31	97.5	4.0	403	US-08-776-246-4	Sequence 4, Appli
32	97.5	4.0	619	US-09-107-532A-4897	Sequence 4897, App
33	96	4.0	331	US-09-724-623-78	Sequence 78, Appli
34	96	4.0	465	US-09-634-238-411	Sequence 411, App
35	96	4.0	693	US-08-463-620-11	Sequence 11, Appli
36	96	4.0	693	US-08-224-917-11	Sequence 11, Appli
37	96	4.0	693	US-08-914-853-11	Sequence 11, Appli
38	96	4.0	693	PCT-US95-03934A-11	Sequence 11, Appli
39	95.5	4.0	340	US-09-789-300A-2	Sequence 2, Appli
40	95.5	4.0	373	US-09-198-452A-416	Sequence 416, App
41	95.5	4.0	406	US-09-134-001C-3544	Sequence 3544, App
42	94.5	3.9	1452	US-09-252-991A-18118	Sequence 18118, A
43	94	3.9	455	US-09-328-352-5839	Sequence 5839, App
44	93	3.8	450	US-09-107-532A-6534	Sequence 6534, App
45	92.5	3.8	440	US-09-107-532A-4071	Sequence 4071, App

ALIGNMENTS

RESULT 1  
US-09-482-273-120  
Sequence 120, Application US/09482273  
Patent No. 6534631  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 71 Human Secreted Proteins  
FILE REFERENCE: P2030P1  
CURRENT APPLICATION NUMBER: US/09/482,273  
EARLIER FILING DATE: 2000-01-13  
EARLIER APPLICATION NUMBER: PCT/US99/15849  
EARLIER FILING DATE: 1999-07-14  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/092,922  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/092,956  
EARLIER FILING DATE: 1998-07-15  
NUMBER OF SEQ ID NOS: 267  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 120  
LENGTH: 473  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURES:  
NAME/KEY: SITE  
LOCATION: (473)  
OTHER INFORMATION: Xaa equals stop translation  
US-09-482-273-120  
Query Match 100.0%; Score 2417; DB 4; Length 473;  
Best Local Similarity 100.0%; Pred. No. 1.6e-232;  
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKFLIFAFGCVHLLSICSGKAICKNGISKRTEEIKERIASCGDVAKAIINLAIVYKAO 60  
DB 1 MKFLIFAFGCVHLLSICSGKAICKNGISKRTEEIKERIASCGDVAKAIINLAIVYKAO 60  
QY 61 NRSYERIALLVDPVGPRLSSKNLEKAIQIMYONLQODGLEKVLHLEVRIPHERGESA 120  
DB 61 NRSYERIALLVDPVGPRLSSKNLEKAIQIMYONLQODGLEKVLHLEVRIPHERGESA 120  
QY 121 VMLSPRIHKAIAIGLSSISGTPPEGTAIVLVTSFDELQRRASBARKIVVYNQPIYNY 180  
DB 121 VMLSPRIHKAIAIGLSSISGTPPEGTAIVLVTSFDELQRRASBARKIVVYNQPIYNY 180  
QY 181 SRTVQVTRTQCAVBAKVGALASLIRSVASFISYSPHTGIEYODGVKPIPTACTIVDAE 240  
DB 181 SRTVQVTRTQCAVBAKVGALASLIRSVASFISYSPHTGIEYODGVKPIPTACTIVDAE 240

QY 241 MMSRASHGKIKIVLQKMGAKTYPDTSFNTVAETGSKYEQVVLVSGHLSMDVGGCA 300  
DB 241 MMSRASHGKIKIVLQKMGAKTYPDTSFNTVAETGSKYEQVVLVSGHLSMDVGGCA 300  
QY 301 MDDGGAFISWEALSLIDGLRPRTTLRLVMTAEQGVGAFOYYQLHKVNISNYSLV 360  
DB 301 MDDGGAFISWEALSLIDGLRPRTTLRLVMTAEQGVGAFOYYQLHKVNISNYSLV 360  
QY 361 MESDAGTFLPTGLQFTGSEKARAMEEWSLLQPLNTQVLSHGSGTINFMIOAGVGA 420  
DB 361 MESDAGTFLPTGLQFTGSEKARAMEEWSLLQPLNTQVLSHGSGTINFMIOAGVGA 420  
QY 421 SLDDLKYKFFPHSHGDTMTVMDPKOMVAAVAVSVYVADMEMLPRS 472  
DB 421 SLDDLKYKFFPHSHGDTMTVMDPKOMVAAVAVSVYVADMEMLPRS 472

RESULT 2  
US-09-079-955-2  
; Sequence 2, Application US/09079955A  
; Patent No. 6465209  
; GENERAL INFORMATION:  
; APPLICANT: Alexander Blinkovsky  
; APPLICANT: Kimberly Brown  
; APPLICANT: Elizabeth Golligltly  
; APPLICANT: Tony Byun  
; APPLICANT: Thomas Mathiasen  
; APPLICANT: Lene V. Kotod  
; APPLICANT: Mikio Fujii  
; APPLICANT: Chigusa Shizuoka  
; TITLE OF INVENTION: Methods For Producing Protein  
; FILE REFERENCE: 5253,500-US  
; CURRENT APPLICATION NUMBER: US/09/079,955A  
; CURRENT FILING DATE: 1998-05-15  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 496  
; TYPE: PRT  
; ORGANISM: Aspergillus oryzae  
US-09-079-955-2

Query Match 9.5%; Score 229.5; DB 4; Length 496;  
Best Local Similarity 22.6%; Pred. No. 4,1e-14;  
Matches 96; Conservative 83; Mismatches 175; Indels 71; Gaps 19;  
QY 77 RLSSGNLEKAIQIMYQNLQODGLEKHLPEVRIPIHMERGESAVMLEPRHKAIALGLG 136  
DB 53 RYFGKADDTVNTLYBELKKTGYDYVKQF-QVHLSMADOTLKVGBEIE-----AK 105  
QY 137 SSIGTPREGITAEVLVVTSPDELQF-RASEARGKIVVYNQ--PYINSTRVYRTQAGAV 192  
DB 106 TMTYSPSEVETADVAVVKNLGCSEADYPSDVEGKVALIKRGCEPFGDXS-----V 155  
QY 193 EAAKGALASLIRSVASFSIYSPHTGIGQYODGVKIPACTIVDAEMSMASHGKI 252  
DB 156 LAARKKAASIVYNNVAGSMAGTLGAQ--SDKGPYSAIVGISLEDDQKLILATAG-SV 212  
QY 253 VIQTMGAKTYPDTSFNTVAETGSKYEQVVLVSGHLSMDVGGAMDDGGAFISWE 312  
DB 213 SYDLVAVDSKQENRT--TYNVVAQTKGSD-PNNVVALGCHTDSVEARGINDSGIISNLV 270  
QY 313 ALSLIKIDGLRPKRTTLRLVMTAEQGVGAFOYY-----QLHKVNI-----SN 356  
DB 271 IAKALTOYSV--KNAVRFLEMTAEERFGLGS--NYVSHLNAATELKNIRLYLNFDMIASPN 327  
QY 357 YSL-VMESDAGTFLPTGLQFTGSEKARAMEEWSLLQPLNTQVLSHGSG-TDINFMIO 414  
DB 328 YALMTYDGGSAFNOSGP--AGSAQIEKLFED--YYSDIDLPHIPTQFGRSDYEAFIL 382  
QY 415 AGVPGASLIDL-----YKFFPHSHGDTMTVMDPKO--MNVAAAVW 455

DB 383 NGIPSGFLTGAEIGMEENASRWGQAGVAYDANYHAAGDMNTNLNHEAFLINSKATAF 442  
QY 456 AVVSY 460  
DB 443 AVATY 447

RESULT 3  
US-09-330-095-1  
; Sequence 1, Application US/09330095  
; Patent No. 6127161  
; GENERAL INFORMATION:  
; APPLICANT: Kikkoman Corporation  
; TITLE OF INVENTION: Leucine Aminopeptidase Gene, Recombinant DNA, and  
; FILE REFERENCE: PH-622  
; CURRENT APPLICATION NUMBER: US/09/330,095  
; CURRENT FILING DATE: 1999-06-11  
; EARLIER APPLICATION NUMBER: JP-164611/1998  
; EARLIER FILING DATE: 1998-06-12  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 481  
; TYPE: PRT  
; ORGANISM: Aspergillus sojae  
US-09-330-095-1

Query Match 9.4%; Score 226.5; DB 3; Length 481;  
Best Local Similarity 22.0%; Pred. No. 7.8e-14;  
Matches 93; Conservative 84; Mismatches 180; Indels 65; Gaps 18;

QY 77 RLSSGNLEKAIQIMYQNLQODGLEKHLPEVRIPIHMERGESAVMLEPRHKAIALGLG 136  
DB 38 RYFGKADDTVNTLYBELKKTGYDYVKQF-QVHLSMADOTLKVGBEIE-----AK 90  
QY 137 SSIGTPREGITAEVLVVTSPDELQF-RASEARGKIVVYNQPYINSTRVYRTQAGAVEA 195  
DB 91 TMTYSPSEVETADVAVVKNLGCSEADYPSDVEGKVALIKRGCEAFG-----DKSVLAA 143  
QY 196 KVGALASLIRSVASFSIYSPHTGIGQYODGVKIPACTIVDAEMSMASHGKIKIVQ 255  
DB 144 KAKAAASIVYNNVAGSMAGTLGAQ--SDKGPYSAIVGISLEDDQKLILAEAG-SVSVD 200  
QY 256 LKMGAKTYPDTSFNTVAETGSKYEQVVLVSGHLSMDVGGAMDDGGAFISWEALS 315  
DB 201 LWDVSKQENRT--TYNVVAQTKGSD-PNNVVALGCHTDSVEARGINDSGIISNLVAK 258  
QY 316 LIKIDGLRPKRTTLRLVMTAEQGVGAFOYY-----QLHKVNI-----SNYSL 359  
DB 259 ALTOYSV--KNAVRFLEMTAEERFGLGS--NYVSHLNAATELKNIRLYLNFDMIASPNVAL 315  
QY 360 VMSDAGTFLPTGLQFTGSEKARAMEEWSLLQPLNTQVLSHGSG-TDINFMIOAGV 417  
DB 316 MTDGDSGAFNOSGP--AGSAQIEKLFED--YYSDIDLPHIPTQFGRSDYEAFILINGI 370  
QY 418 PGASLIDL-----YKFFPHSHGDTMTVMDPKO--MNVAAAVW 458  
DB 371 PAGGLFGAGIGMEENASRWGQAGVAYDANYHAAGDMNTNLNHEAFLINSKATARAFA 430  
QY 459 SY 460  
DB 431 TY 432

RESULT 4  
US-09-252-991A-18292  
; Sequence 18292, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS



CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/973,337A  
FILING DATE: 05 NOV 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 1747/41426  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
TELEX: 422523 COOP UI  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 750 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-325-553-2

Query Match 6.8%; Score 165; DB 1; Length 750;  
Best Local Similarity 19.2%; Pred. No. 2.2e-07;  
Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

QY 3 FLIFAFGCVHLLSLCSGKAIKNGISKRTFEEIKKEIASCGDVAKAIINLAVYGKQNR 62  
DB 32 FFLGFLGFWFKSSNEATNITPRKNNKAFDELKAE-----NIKKFLYNF----- 77  
QY 63 SYERLALLVDIVGPRISGSK--NLEKAIQIYMONLOODGLEKVLHPV-----RIPHW 113  
DB 78 -----TQIPLHAGTEQNFOIAKQIOSQWKEFGDSVLAHYDVLSPYKTHPNY 127  
QY 114 -----ERGEE--SAVMLEPRHIAIKAILGLSSIGTPPEGITAEVLVYTSFDELQRRASEA 166  
DB 128 ISIINEDGELFNSTLSEF-----PPGYENVSDIVPPFSAFSPQGM-P 170  
QY 167 RGIIVVNOPIYNSRT-----VOYRTQGAVEAAKV-----ALASLIRSV 207  
DB 171 EGDLY-----YVNAKTEDEPFLEKRDMDKINCSGKIYIARYGVFRGNKVKNAQLAGAGV 225  
QY 208 ASFS-----IYSPHTGIGQYOD----- 224  
DB 226 ILYSDPADYFAP--GVKSYPDGMNLPGGGVORGNILNLNGADPLTPGYANEVAYRRGI 283  
QY 225 -----GVKPIPTACTIVEDAEMMSRMAHGIRKIVIQKGAITYPD----- 265  
DB 284 AEAVGLPEIPVHPPIGYDDAQKLE-----KKGGSAPPDSSWRGSLKVPYNVGP 331  
QY 266 -----TDSFNTVAEITGSKYEQVVLVSGHLSMDVGGAMDG 304  
DB 332 GFTGNFSTQVKVMHISTNEVTRINVIQTLRGAVEBDRVYILGHRDSWVF--GGIDPQ 389  
QY 305 GGAFISWE--ALSLIKDLGLRPKRTLRVLWMTAEBOGVGAFOYYQLHKVNISVSLVM 361  
DB 390 SGAIVHETIVRSFGTLKKEGWRPRRTILFASWDAEFGILGSTMAE-----ENSRLLQ 443  
QY 362 ESDAGTFLPTGLQFTGSEKARAIMBEVM-SLLQPLNTIYOVLSHGEGTDINFIQAGVGA 420  
DB 444 ERGV-AVINADSSIEGNVTLRVDCITPLMYSLVH--NLTKELKSPD-----EGFEGK 491  
QY 421 SLDDLYK 428  
DB 492 SLYESWTK 499

RESULT 8  
US-08-394-152A-2  
Sequence 2, Application US/08394152A  
Patent No. 5935818  
GENERAL INFORMATION:  
APPLICANT: Israel, Ron S.  
APPLICANT: Heston, Warren D.W.  
APPLICANT: Fair, William R.

TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM 330 466 DX2  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/394,152A  
FILING DATE: 24-FEB-95  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 41426-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 750 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-394-152A-2

Query Match 6.8%; Score 165; DB 2; Length 750;  
Best Local Similarity 19.2%; Pred. No. 2.2e-07;  
Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

QY 3 FLIFAFGCVHLLSLCSGKAIKNGISKRTFEEIKKEIASCGDVAKAIINLAVYGKQNR 62  
DB 32 FFLGFLGFWFKSSNEATNITPRKNNKAFDELKAE-----NIKKFLYNF----- 77  
QY 63 SYERLALLVDIVGPRISGSK--NLEKAIQIYMONLOODGLEKVLHPV-----RIPHW 113  
DB 78 -----TQIHLHAGTEQNFOIAKQIOSQWKEFGDSVLAHYDVLSPYKTHPNY 127  
QY 114 -----ERGEE--SAVMLEPRHIAIKAILGLSSIGTPPEGITAEVLVYTSFDELQRRASEA 166  
DB 128 ISIINEDGELFNSTLSEF-----PPGYENVSDIVPPFSAFSPQGM-P 170  
QY 167 RGIIVVNOPIYNSRT-----VOYRTQGAVEAAKV-----ALASLIRSV 207  
DB 171 EGDLY-----YVNAKTEDEPFLEKRDMDKINCSGKIYIARYGVFRGNKVKNAQLAGAGV 225  
QY 208 ASFS-----IYSPHTGIGQYOD----- 224  
DB 226 ILYSDPADYFAP--GVKSYPDGMNLPGGGVORGNILNLNGADPLTPGYANEVAYRRGI 283  
QY 225 -----GVKPIPTACTIVEDAEMMSRMAHGIRKIVIQKGAITYPD----- 265  
DB 284 AEAVGLPEIPVHPPIGYDDAQKLE-----KKGGSAPPDSSWRGSLKVPYNVGP 331  
QY 266 -----TDSFNTVAEITGSKYEQVVLVSGHLSMDVGGAMDG 304  
DB 332 GFTGNFSTQVKVMHISTNEVTRINVIQTLRGAVEBDRVYILGHRDSWVF--GGIDPQ 389  
QY 305 GGAFISWE--ALSLIKDLGLRPKRTLRVLWMTAEBOGVGAFOYYQLHKVNISVSLVM 361  
DB 390 SGAIVHETIVRSFGTLKKEGWRPRRTILFASWDAEFGILGSTMAE-----ENSRLLQ 443  
QY 362 ESDAGTFLPTGLQFTGSEKARAIMBEVM-SLLQPLNTIYOVLSHGEGTDINFIQAGVGA 420  
DB 444 ERGV-AVINADSSIEGNVTLRVDCITPLMYSLVH--NLTKELKSPD-----EGFEGK 491

QY 421 SLDDLK 428  
Db 492 SLVESWTK 499

## RESULT 9

US-09-044-668-2  
; Sequence 2, Application US/09044668  
; Patent No. 6150508

## GENERAL INFORMATION:

APPLICANT: Murphy, Gerald P.  
APPLICANT: Boynton, Alton L.  
APPLICANT: Holmes, Eric H.  
APPLICANT: Tino, William Thomas  
TITLE OF INVENTION: MONOCLONAL ANTIBODIES SPECIFIC  
TITLE OF INVENTION: FOR THE EXTRACELLULAR DOMAIN OF PROSTATE-SPECIFIC  
TITLE OF INVENTION: MEMBRANE ANTIGEN  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds, LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2811

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/044,668  
FILING DATE: 18-MAR-1998  
CLASSIFICATION: 536  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER:

## ATTORNEY/AGENT INFORMATION:

NAME: Baldwin, Geraldine F  
REGISTRATION NUMBER: 31,232  
REFERENCE/DOCKET NUMBER: 8511-0013-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-493-4935  
TELEFAX: 650-493-5556  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 750 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: No. 6150508e  
US-09-044-668-2

Query Match 6.8%; Score 165; DB 3; Length 750;  
Best Local Similarity 19.2%; Pred. No. 2.2e-07;  
Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

QY 3 FLIFAFPGVHLTLSCGKAIKNGISKRTFEEIKERIASCGVAKAIINLAVYKAQR 62  
Db 32 FFLGLFPGWFIKSNATNTITPKHMKAFDLKAE-----NKKFLVNF----- 77  
QY 63 SYERLALLVDTVGRISGSK--NLEKAIQIMYQNTLOODGLEKVNHEPV-----RIPHW 113  
Db 78 -----TQIPHLAGTEQNFOIAKQIOSWKKEFGIDSVELAHYVDVLLSYPKTKHPNY 127  
QY 114 -----ERGER--SAVMLEPRIRHKIALILGLSSIGTPREGITAEVLVYTSFDELQRASEA 166  
Db 128 ISINEDNEIFNTSLFEP-----PPGYNVSDIVPPSAFSPQGM-P 170  
QY 167 RGIIVVNOPYINVSRT-----VOYRTOGAVEAKVG-----ALASLIRSV 207

Db 171 EGBLV-----YVYVARTEDFFKLERDMKINCSKIVARIYKVFGRNKNVQAOLAGAKGV 225  
QY 208 ASFS-----IYSPHTGIOEYOD----- 224  
Db 226 ILVSDPADYFAP--GVKSYDPDGMNLPGGGVORGNIILNLGAGDPLPPGYANAYARRGI 283  
QY 225 -----GVKPIPTACITVEDAEEMSRMASHGIKIYIOLMGAKTYPD----- 265  
Db 284 ABAVGLPSIVVHPHIGYDAQKLE-----KMGSAAPPDSSWRSLKVPYVNGP 331  
QY 266 -----TDFNTVAITTSKPIPEQVVLVSGHLDSDVQOGAMDG 304  
Db 332 GFTGNFSTQKXKHISTNEVTRIVYVIGTLRGAVERDRYVILGHRDSWVF--GCIDPQ 389  
QY 305 GGAFLSWE---ALSLIKDGLRPRKTRLVLTAEQGGGAGFOYQOLHKVINSYSLVM 361  
Db 390 SGAIVVHEIVRSFGTLKESGMRPRRTILFASWPAEEGLGISTEWE-----ENSRLLQ 443  
QY 362 ESDAGFLPTGLOFTGSEKARAIMEEVM--SLQPLNITQVLSHGEGTDINFWIQAGVPGA 420  
Db 444 ERGV-AVINADSSIEGNYTLRVDCITPLMYSLVH--NUTKELKSPD-----EGREGK 491  
QY 421 SLDDLK 428  
Db 492 SLVESWTK 499

## RESULT 10

US-08-705-477E-2  
; Sequence 2, Application US/08705477E  
; Patent No. 6569432

## GENERAL INFORMATION:

APPLICANT: Israeli, Ron S  
APPLICANT: Heston, Warren D.W.  
APPLICANT: Fair, William R.  
APPLICANT: Overfelli, Ouathek  
APPLICANT: Pinto, John

TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND USES THEREOF

FILE REFERENCE: 1769/41426-G

CURRENT APPLICATION NUMBER: US/08/705,477E

CURRENT FILING DATE: 1996-08-29

NUMBER OF SEQ ID NOS: 128

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 750

TYPE: PRT

ORGANISM: Homo sapiens

US-08-705-477E-2

Query Match 6.8%; Score 165; DB 4; Length 750;  
Best Local Similarity 19.2%; Pred. No. 2.2e-07;  
Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

QY 3 FLIFAFPGVHLTLSCGKAIKNGISKRTFEEIKERIASCGVAKAIINLAVYKAQR 62  
Db 32 FFLGLFPGWFIKSNATNTITPKHMKAFDLKAE-----NKKFLVNF----- 77  
QY 63 SYERLALLVDTVGRISGSK--NLEKAIQIMYQNTLOODGLEKVNHEPV-----RIPHW 113  
Db 78 -----TQIPHLAGTEQNFOIAKQIOSWKKEFGIDSVELAHYVDVLLSYPKTKHPNY 127  
QY 114 -----ERGER--SAVMLEPRIRHKIALILGLSSIGTPREGITAEVLVYTSFDELQRASEA 166  
Db 128 ISINEDNEIFNTSLFEP-----PPGYNVSDIVPPSAFSPQGM-P 170  
QY 167 RGIIVVNOPYINVSRT-----VOYRTOGAVEAKVG-----ALASLIRSV 207  
Db 171 EGBLV-----YVYVARTEDFFKLERDMKINCSKIVARIYKVFGRNKNVQAOLAGAKGV 225  
QY 208 ASFS-----IYSPHTGIOEYOD----- 224  
Db 226 ILVSDPADYFAP--GVKSYDPDGMNLPGGGVORGNIILNLGAGDPLPPGYANAYARRGI 283



```

Db      195 -GVASYPDGMNLPGGGVORGNILNLNGADPLTPGPANEYAVRGLAEAVGLPSIPVHP 253
Qy      234 ITVDAEMSMASHGKIYIOLKMGATYPD----- 265
Db      254 IGYVDAQKLE-----KMGSAAPPDSSWRGSLKVPYVNGPFTGNFSTQKYKM 301
Qy      266 -----TDSFNTVAEITGSKYPEQVYLVSGHLSMDVQAGAMDGGAFISWE---AL 314
Db      302 HIHSTNEVTRIVYVIGLRGAVEPDRYVILGHRDSWVF--GGIDPQGAHVHEIVRSF 359
Qy      315 SLIDGLRPKRTLRVLTAEBOGVGAFQYQYOLHKVNISNYSLVMSDAGTFLPTGLQ 374
Db      360 GTLKKEGMRPRRTILFASMDAEFGLGSTEMAE-----ENSRLLQERGVAAYINADSS 412
Qy      375 FTGSEKARAMEVW-SILOPLNITOVLSHGEQTDINFWIOAGVPASLLDDLYK 428
Db      413 IEGVYTRVDCPTPLMSLVH--NLTKEKSPD-----EGREGSLYESWTK 456

```

## RESULT 13

```

PCT-US95-11720-18
; Sequence 18, Application PC/TUS9511720
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: EXPRESSION AND EXPORT TECHNOLOGY OF
; TITLE OF INVENTION: PROTEINS AS IMMUNOPROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
; ADDRESSEE: THIBEAULT
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/11720
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, EDMUND R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: FIP-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7000
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 707 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..707
; OTHER INFORMATION: /note="EXTRACELLULAR DOMAIN OF
; OTHER INFORMATION: PSMA"
; PCT-US95-11720-18

```

Query Match 6.6%; Score 160; DB 5; Length 707;

Best Local Similarity 19.6%; Pred No. 6.4e-07; Matches 93; Conservative 60; Mismatches 144; Indels 178; Gaps 21;

```

Qy      76 PRLSGSK--NLKXAIQIWMYQNLQDGLKVKLEPV-----RIPHW-----ERGER--S 119
Db      38 PRLAGTEQWFLAKQIQSQWKEFGDLSDVELAHYDVLLSYPNKTHPNYISIIINEDGNELFN 97

```

```

Qy      120 AVMLEPRHIAILGLSSIGTPPEGITAEVLVVTSEFDELQRRASEARGIWVYNOPYIN 179
Db      98 TSLFEP-----PPPGYENSDIVPEPSASPSQGM-PEGGLV-----YVN 135
Qy      180 YSRT-----VOYRTQGAVEAAKVG-----ALASHIRSVASFS---IYSPH 216
Db      136 YARTEDEPKLRDMKINCSKIVARYGKVRGNKVNQAOLAGAKGVIYLSDDADYFAP- 194
Qy      217 TGIQEOYD-----GVPKIPTAC 233
Db      195 -GVASYPDGMNLPGGGVORGNILNLNGADPLTPGPANEYAVRGLAEAVGLPSIPVHP 253
Qy      234 ITVDAEMSMASHGKIYIOLKMGATYPD----- 265
Db      254 IGYVDAQKLE-----KMGSAAPPDSSWRGSLKVPYVNGPFTGNFSTQKYKM 301
Qy      266 -----TDSFNTVAEITGSKYPEQVYLVSGHLSMDVQAGAMDGGAFISWE---AL 314
Db      302 HIHSTNEVTRIVYVIGLRGAVEPDRYVILGHRDSWVF--GGIDPQGAHVHEIVRSF 359
Qy      315 SLIDGLRPKRTLRVLTAEBOGVGAFQYQYOLHKVNISNYSLVMSDAGTFLPTGLQ 374
Db      360 GTLKKEGMRPRRTILFASMDAEFGLGSTEMAE-----ENSRLLQERGVAAYINADSS 412
Qy      375 FTGSEKARAMEVW-SILOPLNITOVLSHGEQTDINFWIOAGVPASLLDDLYK 428
Db      413 IEGVYTRVDCPTPLMSLVH--NLTKEKSPD-----EGREGSLYESWTK 456

```

## RESULT 14

```

US-08-547-197-1
; Sequence 1, Application US/08547197
; Patent No. 5691157
; GENERAL INFORMATION:
; APPLICANT: Gong, Joseph K.
; APPLICANT: Glomski, Chester A.
; TITLE OF INVENTION: A METHOD FOR DETECTING A MAMMAL'S PRIOR
; TITLE OF INVENTION: EXPOSURE TO RADIATION OR RADIOMETRIC AGENTS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/547,197
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rogalsky, Peter
; REGISTRATION NUMBER: 38,601
; REFERENCE/DOCKET NUMBER: 19226/580 (R-5228)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1634
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 622 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-547-197-1

```

Query Match 5.9%; Score 143.5; DB 1; Length 622;





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:47:05 ; Search time 39.3941 Seconds  
(without alignments)  
2237.645 Million cell updates/sec

Title: US-09-745-763-36  
Perfect score: 2417  
Sequence: 1 MKFLIFAFGCVHLSICSG.....AWNAVSVVADMEMLPRS 472

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 696363 seqs, 186758610 residues  
Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubppaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubppaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubppaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubppaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubppaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2417	100.0	472	9	US-09-745-763-36
2	2417	100.0	473	11	US-09-984-271-120
3	188	7.8	740	12	US-10-325-430-15
4	165	6.8	33	11	US-09-984-271-224
5	165	6.8	34	11	US-09-984-271-187
6	165	6.8	750	9	US-09-822-827-944
7	165	6.8	750	10	US-09-895-793-944
8	165	6.8	750	12	US-10-341-434-71
9	165	6.8	750	15	US-10-117-937-4
10	165	6.8	750	15	US-10-094-699-1
11	165	6.8	750	15	US-10-205-823-132
12	163	6.7	750	10	US-09-978-295A-618
13	163	6.7	750	10	US-09-978-697-618
14	163	6.7	750	10	US-09-978-192A-618
15	163	6.7	750	10	US-09-999-832A-618

16	163	6.7	750	11	US-09-978-189-618	Sequence 618, App
17	163	6.7	750	11	US-09-978-608A-618	Sequence 618, App
18	163	6.7	750	11	US-09-978-585A-618	Sequence 618, App
19	163	6.7	750	11	US-09-978-191A-618	Sequence 618, App
20	163	6.7	750	11	US-09-978-403A-618	Sequence 618, App
21	163	6.7	750	11	US-09-978-564A-618	Sequence 618, App
22	163	6.7	750	11	US-09-999-833A-618	Sequence 618, App
23	163	6.7	750	11	US-09-981-915A-618	Sequence 618, App
24	163	6.7	750	11	US-09-978-824-618	Sequence 618, App
25	163	6.7	750	11	US-09-918-585A-618	Sequence 618, App
26	163	6.7	750	11	US-09-978-423A-618	Sequence 618, App
27	163	6.7	750	11	US-09-978-193A-618	Sequence 618, App
28	163	6.7	750	11	US-09-999-830A-618	Sequence 618, App
29	163	6.7	750	11	US-09-978-757A-618	Sequence 618, App
30	163	6.7	750	11	US-09-978-187B-618	Sequence 618, App
31	163	6.7	750	11	US-09-978-643A-618	Sequence 618, App
32	163	6.7	750	12	US-09-978-375A-618	Sequence 618, App
33	163	6.7	750	12	US-09-978-188A-618	Sequence 618, App
34	163	6.7	750	12	US-09-978-298A-618	Sequence 618, App
35	163	6.7	750	12	US-10-143-031A-618	Sequence 618, App
36	163	6.7	750	12	US-10-002-967A-618	Sequence 618, App
37	163	6.7	750	12	US-10-017-083A-618	Sequence 618, App
38	163	6.7	750	12	US-10-143-030A-618	Sequence 618, App
39	163	6.7	750	12	US-10-199-672-104	Sequence 104, App
40	163	6.7	750	12	US-10-187-749-104	Sequence 104, App
41	163	6.7	750	12	US-10-194-457-104	Sequence 104, App
42	163	6.7	750	12	US-10-145-128A-618	Sequence 618, App
43	163	6.7	750	12	US-10-184-642-104	Sequence 104, App
44	163	6.7	750	12	US-10-196-747-104	Sequence 104, App
45	163	6.7	750	12	US-10-173-689-104	Sequence 104, App

ALIGNMENTS

RESULT 1  
US-09-745-763-36  
Sequence 36, Application US/09745763  
Patent No. US20020065394A1  
GENERAL INFORMATION:  
APPLICANT: McCoy, Kenneth  
McCoy, John M.  
Lavallee, Edward R.  
Collins-Racie, Lisa A.  
Evans, Cheryl  
Merberg, David  
Treacy, Maurice  
Spaulding, Vikki  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM  
NUMBER OF SEQUENCES: 219  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/745,763  
FILING DATE: 18-Jun-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Sprunger, Suzanne A.  
REGISTRATION NUMBER: 41,323  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8284  
TELEFAX: (617) 876-5851



```

Db      56 KAENIKSFLRSFTKL-----PHLAGTEQNFLAKKIOTQWKGFLDSAKLVHYDVL 107
Qy      109 RIPHMERGESAWMLBPRIRKIALGLGSSIGTPEGTALVLTSTGFDLQRRASBARG 168
Db      108 SYRETEANNTIYSIDE---HETELFKT-SYLEPPDDGENTNIVPYNNAFSAQGM-PEG 162
Qy      169 KIIVYNOPYINTSYRTVOY-----RTQGAVEAAKVAL-----ASLIRSVASFSI 212
Db      163 DLV-----YVNVARTDEPFKLEREMGINCTGKIYIARXGKIFRGNKYNMMLAGAIIGIL 217
Qy      213 YS-----PHTGIO-----EYOD 224
Db      218 YSPBADYFAEVOYPRKGMNLPGTAQRGNVNLNGADPLTPGYPAKEYTFRLDVEGV 277
Qy      225 GVPIKIPACTIVEDAEMMSRMAHSGIKIVIQKMGAKT-----YPTDTSF----- 269
Db      278 GIPRIPIPHICYNDAEILRLTG-GIAPPDKSMGALNVSTISPGTGSDFPKVMHV 336
Qy      270 -----NTVAETGSKYPEQVVLVSGHLSMDVQOGAMDGGAFISWE---ALSL 316
Db      337 YNINKIRIVNVGTIGTSVEPRRYVLGHRDSWVF--GAIPTSGAVLQELIARSPGK 394
Qy      317 IKDLGLRPKRTLRLVLTAEBOGVGAFOYQLHKVNISYSLVMSDAGTFELFTGLQFT 376
Db      395 LMSKMPRRRTIIPASWDABEFGLGSTEWAE-----ENVKILQERSI-AVINSDSIE 447
Qy      377 GSEKARA-----IMEEWSLLOPL-----NITQVLSHGE 405
Db      448 GNYTLRVDCPTPLLYQLVYKLTKEIPSPDDGESKSLYESWLEKPPSPENKLPRLINKGS 507
Qy      406 GTDINFMIQ-AGVGA-----SLDDLKYKFFHHSHGDTMTVM---DP---KOMNVA 451
Db      508 GSDEAATFORGLGASGARATYKNNKKTDXSYRYHTIYEFELVEKFDYPTFKQLSLVA 567
Qy      452 AAWAVVSYVVADEMLP 470
Db      568 QLRGALV-YELVD-SKIIP 584

```

```

RESULT 4
US-09-984-271-224
; Sequence 224, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984, 271
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482, 273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092, 921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092, 922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092, 956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 224
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-271-224

```

```

Query Match      6.8%; Score 165; DB 11; Length 33;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      440 MTWMDPKOMNVAAVAVVSYVVADEMLPRS 472
|||||

```

```

Db      1 MTWMDPKOMNVAAVAVVSYVVADEMLPRS 33
RESULT 5
US-09-984-271-187
; Sequence 187, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/984, 271
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/482, 273
; PRIOR FILING DATE: 2000-01-13
; PRIOR APPLICATION NUMBER: PCT/US99/15849
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: 60/092, 921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092, 922
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/092, 956
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 187
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals stop translation
US-09-984-271-187

```

```

Query Match      6.8%; Score 165; DB 11; Length 34;
Best Local Similarity 100.0%; Pred. No. 3.3e-09;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      440 MTWMDPKOMNVAAVAVVSYVVADEMLPRS 472
|||||

```

```

Db      1 MTWMDPKOMNVAAVAVVSYVVADEMLPRS 33
RESULT 6
US-09-822-827-944
; Sequence 944, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822, 827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: PasteSeq for Windows Version 3.0
; SEQ ID NO 944
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-822-827-944

```

```

Query Match      6.8%; Score 165; DB 9; Length 750;
Best Local Similarity 19.2%; Pred. No. 5e-07;
Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;
Qy      3 FLIAPFGVYHLISLGSKXKIGISKRPFEIKERIASGDAKAIINLAAYGKAQNR 62
|||||
Db      32 FFLGLFLGFWFISSNATITTPAHNKAFLDELKAE-----NIKKLVNF----- 77
|||||
Qy      63 SYERLALVDTVGPRLSGSK---NLEKAIQIMYQNTLODGLKRVHLBPV-----RIPHW 113
|||||

```

```

Db      78 -----TQIPLHAGTEQNFOLAKQIOSQWKEFGLDSVELAHYDVLSTYPNKTHPNY 127
Qy      114 -----ERGBE--SAVMELEPRHKAIIILGSSICTPPREGITAEVLVYTSFDELQGRASEA 166
Db      128 ISIINEDNEIENFTSLFSP-----PPPEYENVSDIVPPFSAFSPQGM-P 170
Qy      167 RGIIVVYNOPIYINSRT-----VOYRTQGAWEAAKV-----ALASLIRSV 207
Db      171 EGDLY-----YVYAKTBDFPKLERDMKINCSGKIYIARQKVPFGNKVKQAOLAGAKGV 225
Qy      208 ASFS-----TSPHTGIGQYOD----- 224
Db      226 ILVSPDADYFAP--GVKSYPDGMNLPGGGVORGNILNLNGAGDPLTPGPANAYARRGI 283
Qy      225 -----GVKPIPTACTIVEDAEMSRMASHGIKIYIOLMGAKTYPD----- 265
Db      284 AEAVGLPSIPVHPIDYIDAQKLE-----KMGSGAPDSSWRGSLKVPYVGP 331
Qy      266 -----TDSFNTVAEITGSKYPEQVVLVSGHLDSDVGGAMDG 304
Db      332 GFTGNFSTQKVKMHIHSTNEVTRINVTIGTLRGAVEPRYVILGHRDSWF--GIDPQ 389
Qy      305 GGAFISWE--ALSLIKDLGRPKRTLRLVMTAEEOGVGAFOYYQLHKVNISYSLW 361
Db      390 SGAIVVHEIYVSFGLTKKEGWRPRTLIFASWDAEEFGLGSTEWAE-----ENSRLQ 443
Qy      362 ESDAGTFLPTGLQTFGSEKARAMEEW-SILOPLNTQVLSHGEGTDINFIQAGVGA 420
Db      444 ERGV-AYINADSIIEGNYTLRVDCPTPLMYSLVH--NLTKELKSPD-----EGFEKG 491
Qy      421 SLDDLYK 428
Db      492 SLYESWTK 499

```

# RESULT 7

US-09-895-793-944

Sequence 944, Application US/09895793

Publication No. US20020192763A1

## GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
 APPLICANT: Dillon, Davin C.  
 APPLICANT: Mitcham, Jennifer L.  
 APPLICANT: Harlocker, Susan L.  
 APPLICANT: Jiang, Yugu  
 APPLICANT: Kalos, Michael D.  
 APPLICANT: Retter, Marc W.  
 APPLICANT: Stolk, John A.  
 APPLICANT: Day, Craig H.  
 APPLICANT: Vedrick, Thomas S.  
 APPLICANT: Carter, Darriek  
 APPLICANT: Li, Samuel X.  
 APPLICANT: Wang, Aijun  
 APPLICANT: Skeiky, Yaelir A.W.  
 APPLICANT: Hepler, William T.  
 APPLICANT: Henderson, Robert A.  
 APPLICANT: Hurai, John  
 APPLICANT: McNeill, Patricia D.  
 APPLICANT: Houghton, Raymond L.  
 APPLICANT: Vinals de Basbols, Carlota  
 APPLICANT: Foy, Teresa  
 APPLICANT: Panger, Gary R.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
 FILE REFERENCE: 210121.534C2  
 CURRENT APPLICATION NUMBER: US/09/895,793  
 CURRENT FILING DATE: 2001-06-29  
 NUMBER OF SEQ ID NOS: 982  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 944  
 LENGTH: 750  
 TYPE: PRT  
 ORGANISM: Homo sapiens

US-09-895-793-944

Query Match 6.8%; Score 165; DB 10; Length 750;

Best Local Similarity 19.2%; Pred. No. 5e-07; Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

```

Qy      3 FLIAPFGVYLLSLGSKAIKNGISKRTFEEIKERIASGCVAKAIINLAVYGAQNR 62
Db      32 FFLGLFEGMWIKSNENATNITPRGNMKAPLDELAKA-----NKKFLYNE----- 77
Qy      63 SYERLALLVDTGVRLSGSK--NLEKAIQIMYONLOODGIEKYNLEPV-----RIPHW 113
Db      78 -----TQIPLHAGTEQNFOLAKQIOSQWKEFGLDSVELAHYDVLSTYPNKTHPNY 127
Qy      114 -----ERGBE--SAVMELEPRHKAIIILGSSICTPPREGITAEVLVYTSFDELQGRASEA 166
Db      128 ISIINEDNEIENFTSLFSP-----PPPEYENVSDIVPPFSAFSPQGM-P 170
Qy      167 RGIIVVYNOPIYINSRT-----VOYRTQGAWEAAKV-----ALASLIRSV 207
Db      171 EGDLY-----YVYAKTBDFPKLERDMKINCSGKIYIARQKVPFGNKVKQAOLAGAKGV 225
Qy      208 ASFS-----TSPHTGIGQYOD----- 224
Db      226 ILVSPDADYFAP--GVKSYPDGMNLPGGGVORGNILNLNGAGDPLTPGPANAYARRGI 283
Qy      225 -----GVKPIPTACTIVEDAEMSRMASHGIKIYIOLMGAKTYPD----- 265
Db      284 AEAVGLPSIPVHPIDYIDAQKLE-----KMGSGAPDSSWRGSLKVPYVGP 331
Qy      266 -----TDSFNTVAEITGSKYPEQVVLVSGHLDSDVGGAMDG 304
Db      332 GFTGNFSTQKVKMHIHSTNEVTRINVTIGTLRGAVEPRYVILGHRDSWF--GIDPQ 389
Qy      305 GGAFISWE--ALSLIKDLGRPKRTLRLVMTAEEOGVGAFOYYQLHKVNISYSLW 361
Db      390 SGAIVVHEIYVSFGLTKKEGWRPRTLIFASWDAEEFGLGSTEWAE-----ENSRLQ 443
Qy      362 ESDAGTFLPTGLQTFGSEKARAMEEW-SILOPLNTQVLSHGEGTDINFIQAGVGA 420
Db      444 ERGV-AYINADSIIEGNYTLRVDCPTPLMYSLVH--NLTKELKSPD-----EGFEKG 491
Qy      421 SLDDLYK 428
Db      492 SLYESWTK 499

```

# RESULT 8

US-10-341-434-71

Sequence 71, Application US/10341434

Publication No. US20030215835A1

## GENERAL INFORMATION:

APPLICANT: Origene Technologies  
 TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes  
 FILE REFERENCE: 9U 204 205 R1  
 CURRENT APPLICATION NUMBER: US/10/341,434  
 CURRENT FILING DATE: 2003-07-18  
 PRIOR APPLICATION NUMBER: US 60/346,164  
 PRIOR FILING DATE: 2002-01-15  
 PRIOR APPLICATION NUMBER: US 60/348,119  
 PRIOR FILING DATE: 2002-01-15  
 NUMBER OF SEQ ID NOS: 238  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 71  
 LENGTH: 750  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-341-434-71

Query Match 6.8%; Score 165; DB 12; Length 750;  
 Best Local Similarity 19.2%; Pred. No. 5e-07;  
 Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

QY 3 FLIAPFGVHLLSLCSGKAI CKNGISKRTFEEIKKEIASCGDVAKAI INLAAYGKAQNR 62  
 DB 32 FFLGLFGLFWFISKSNEATNITPRKHNKAFLELDEKAE-----NKKFLVNF----- 77  
 QY 63 SYERLALLVDTVGRRLSGSK---NLEKAIQIMYQNLQODGLEKVNLEPV-----RIPHW 113  
 DB 78 -----TQIPLHLAGTEQNFOLAKQIOSQWKEFGJDSVELAHYVLLSYPNKTPHNY 127  
 QY 114 -----ERGEE--SAVMLEPRHKAIALGLSSIGTPEGITAEVLVVTSPDELQRRASEA 166  
 DB 128 ISIIINEGNEIFNTSLPEP-----PPGGENVSIDIVPPSAPSPQGM-P 170  
 QY 167 RGIIVVNPQYINYSRT-----VOYRTOGAVEAKVG-----ALASLIRSV 207  
 DB 171 EGDLY-----YVYARTEDEPFKLERDKINCSKIYARVGKVRGKVKNAQLAGAKGV 225  
 QY 208 ASFS-----IYSPHTGIOEYOD----- 224  
 DB 226 ILVSDPADYFAP--GVKSYPDGMNLPGGGVQGRNIIINLNGADPLTPGYPANEBAYARRGI 283  
 QY 225 -----GVKPIPTACTITVEDAEMMSHAGIKIYIOLKMGAKTYPD----- 265  
 DB 284 AEAVGLPSIIVHPIGYIDAOQLLE-----KMGGSAPDSSWRGSLKVPYVGP 331  
 QY 266 -----TDSFNTVAEITGSKYPEQVVLVSGHLDSDMDVQCAMDDG 304  
 DB 332 GFTGNFSTOKVKMHISTNEVTIRIYVIGTLRGAVERDRYVILGHRDSDWF--GGIDPQ 389  
 QY 305 GGAFTSWE--ALSILKDLGLRPKRTLRLVMTAEBOGCVGAFYQOLHKVNISNYSLVW 361  
 DB 390 SGAIVHVEIYRSFGLTKKEGWRPRRTILFASWDAEEFGLLGSTEMAE-----ENSRLQ 443  
 QY 362 ESDAGTFLPGLOFTGSEKRAIMEBYW--SLLOPLNTQVLSHSEGDINDFWIOAGVPGA 420  
 DB 444 ERGV-AVINADSSIEGNTTLRVDCPTPLMYSLVH--NLTKELKSPD-----EGFEGK 491  
 QY 421 SLDDLYK 428  
 DB 492 SLVESWTK 499

RESULT 9  
 US-10-117-937-4  
 ; Sequence 4, Application US/10117937  
 ; Publication No. US20030220239A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CTL IMMUNO THERAPIES CORP.  
 ; APPLICANT: SIMARD, John, J.L.  
 ; APPLICANT: DIAMOND, David, C.  
 ; APPLICANT: Liu, Liping  
 ; APPLICANT: Xie, Zhidong  
 ; TITLE OF INVENTION: EPTOPE SEQUENCES  
 ; FILE REFERENCE: CTLIMM.027A  
 ; CURRENT APPLICATION NUMBER: US/10/117, 937  
 ; CURRENT FILING DATE: 2002-04-04  
 ; PRIOR APPLICATION NUMBER: US 60/282,211  
 ; PRIOR FILING DATE: 2001-04-06  
 ; PRIOR APPLICATION NUMBER: US 60/337,017  
 ; PRIOR FILING DATE: 2001-11-07  
 ; PRIOR APPLICATION NUMBER: US 60/363,210  
 ; PRIOR FILING DATE: 2002-03-07  
 ; NUMBER OF SEQ ID NOS: 602  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 4  
 ; LENGTH: 750  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-117-937-4

Query Match 6.8%; Score 165; DB 12; Length 750;  
 Best Local Similarity 19.2%; Pred. No. 5e-07;  
 Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

QY 3 FLIAPFGVHLLSLCSGKAI CKNGISKRTFEEIKKEIASCGDVAKAI INLAAYGKAQNR 62  
 DB 32 FFLGLFGLFWFISKSNEATNITPRKHNKAFLELDEKAE-----NKKFLVNF----- 77  
 QY 63 SYERLALLVDTVGRRLSGSK---NLEKAIQIMYQNLQODGLEKVNLEPV-----RIPHW 113  
 DB 78 -----TQIPLHLAGTEQNFOLAKQIOSQWKEFGJDSVELAHYVLLSYPNKTPHNY 127  
 QY 114 -----ERGEE--SAVMLEPRHKAIALGLSSIGTPEGITAEVLVVTSPDELQRRASEA 166  
 DB 128 ISIIINEGNEIFNTSLPEP-----PPGGENVSIDIVPPSAPSPQGM-P 170  
 QY 167 RGIIVVNPQYINYSRT-----VOYRTOGAVEAKVG-----ALASLIRSV 207  
 DB 171 EGDLY-----YVYARTEDEPFKLERDKINCSKIYARVGKVRGKVKNAQLAGAKGV 225  
 QY 208 ASFS-----IYSPHTGIOEYOD----- 224  
 DB 226 ILVSDPADYFAP--GVKSYPDGMNLPGGGVQGRNIIINLNGADPLTPGYPANEBAYARRGI 283  
 QY 225 -----GVKPIPTACTITVEDAEMMSHAGIKIYIOLKMGAKTYPD----- 265  
 DB 284 AEAVGLPSIIVHPIGYIDAOQLLE-----KMGGSAPDSSWRGSLKVPYVGP 331  
 QY 266 -----TDSFNTVAEITGSKYPEQVVLVSGHLDSDMDVQCAMDDG 304  
 DB 332 GFTGNFSTOKVKMHISTNEVTIRIYVIGTLRGAVERDRYVILGHRDSDWF--GGIDPQ 389  
 QY 305 GGAFTSWE--ALSILKDLGLRPKRTLRLVMTAEBOGCVGAFYQOLHKVNISNYSLVW 361  
 DB 390 SGAIVHVEIYRSFGLTKKEGWRPRRTILFASWDAEEFGLLGSTEMAE-----ENSRLQ 443  
 QY 362 ESDAGTFLPGLOFTGSEKRAIMEBYW--SLLOPLNTQVLSHSEGDINDFWIOAGVPGA 420  
 DB 444 ERGV-AVINADSSIEGNTTLRVDCPTPLMYSLVH--NLTKELKSPD-----EGFEGK 491  
 QY 421 SLDDLYK 428  
 DB 492 SLVESWTK 499

RESULT 10  
 US-10-094-699-1  
 ; Sequence 1, Application US/10094699  
 ; Publication No. US20030046714A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SIMARD, John, J.L.  
 ; APPLICANT: DIAMOND, David, C.  
 ; TITLE OF INVENTION: ANTI-NEOVASCULATURE PREPARATIONS FOR  
 ; TITLE OF INVENTION: CANCER  
 ; FILE REFERENCE: CTLIMM.015A  
 ; CURRENT APPLICATION NUMBER: US/10/094,699  
 ; CURRENT FILING DATE: 2002-06-21  
 ; PRIOR APPLICATION NUMBER: US 60/274,063  
 ; PRIOR FILING DATE: 2001-03-07  
 ; NUMBER OF SEQ ID NOS: 109  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 750  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-10-094-699-1

Query Match 6.8%; Score 165; DB 15; Length 750;  
 Best Local Similarity 19.2%; Pred. No. 5e-07;  
 Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

QY 3 FLIAPFGVHLLSLCSGKAI CKNGISKRTFEEIKKEIASCGDVAKAI INLAAYGKAQNR 62  
 DB 32 FFLGLFGLFWFISKSNEATNITPRKHNKAFLELDEKAE-----NKKFLVNF----- 77  
 QY 63 SYERLALLVDTVGRRLSGSK---NLEKAIQIMYQNLQODGLEKVNLEPV-----RIPHW 113

Db 78 -----TQIPLHAGTEQNFOIAKOIOSQWKEFGDLSVELAHYDVLLSYPNKTHPNY 127  
Qy 114 -----ERGEE--SAVMEPRRIHKIAIILGSSICTPPEGITAEVLYVTSFDELQRRASEA 166  
Db 128 ISIINEDENETFNISLFEF-----PPPEYENVSIVPPFSAFSPQGM-P 170  
Qy 167 RGIIVVYNQPIYINSRT-----VOYRTOGAWEAAKV-----ALASLIRSV 207  
Db 171 EGDLY-----YVNAKTEDPFKLERDMKINCSCGIIVARIGKVRGNKYNQAOLAGAKGV 225  
Qy 208 ASFS-----IYSPHTGIEYOD----- 224  
Db 226 ILYSDPADYFAP--GVKSYDPDGMNLPGGCVQGRNIIILNAGADPLTPGPANEYAYRRGI 283  
Qy 225 -----GVKPIPTACTIVEDAEMSRMASHGIKIYIQLKMGAKTYPD----- 265  
Db 284 AEAVGLPSIPVHPPIGYDDAKLLE-----KMGGSAPPDSSWRGSLKVPYNVGP 331  
Qy 266 -----TDSFNTVAETGSKYPEQVVLVSGHLDSDMDVGQAGMDG 304  
Db 332 GFTGNFSTQVKYKMHISTNEVTRIVNIGTLRGAVEPRRYVILGHRDSWVF--GGIDPQ 389  
Qy 305 GGAFISWE--ALSILKDLGLRPKRTLRVLVMTAEQGVGAFOYYQLHKVNISNYSLVW 361  
Db 390 SGAIVVHEIVRSFGLTKKEGWRPRTILFASWDAEEFGLGSTEWAE-----ENSRLLQ 443  
Qy 362 ESDAGTFLPTGLQTFGSEKARAMEEW-SLLQPLNITOVLSHGCTDINFWIQAIGVGA 420  
Db 444 ERGV-AYINADSSIEGNYTLRVDCPTPLMYSLVH--NLTKELKSPD-----EGFEGK 491  
Qy 421 SLDDLYK 428  
Db 492 SLYESWTK 499

RESULT 11  
US-10-205-823-132  
; Sequence 132, Application US/10205823  
; Publication No. US20030108963A1  
GENERAL INFORMATION:  
APPLICANT: Schlegel, Robert  
APPLICANT: Monahan, John E.  
APPLICANT: Endegge, Wilson O.  
APPLICANT: Gannavarapu, Manjula  
APPLICANT: Gothatcheva, Bella  
APPLICANT: Hoerschen, Sebastian  
APPLICANT: Kamatkar, Shubhangi  
APPLICANT: Womsey, Angela M.  
APPLICANT: Glatc, Karen  
APPLICANT: Zhao, Xumei  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
FILE REFERENCE: MRI-044  
CURRENT APPLICATION NUMBER: US/10/205,823  
CURRENT FILING DATE: 2002-07-25  
PRIOR APPLICATION NUMBER: 60/307,982  
PRIOR FILING DATE: 2001-07-25  
PRIOR APPLICATION NUMBER: 60/314,356  
PRIOR FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/325,020  
PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: 60/341,746  
PRIOR FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: 60/362,158  
PRIOR FILING DATE: 2002-03-05  
NUMBER OF SEQ ID NOS: 455  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 132  
LENGTH: 750  
TYPE: PRT  
ORGANISM: Homo sapiens

US-10-205-823-132

Query Match 6.8%; Score 165; DB 15; Length 750;  
Best Local Similarity 19.28; Pred. No. 5e-07;  
Matches 105; Conservative 67; Mismatches 114; Indels 202; Gaps 23;

Qy 3 FLIAPFGVYLLSLGSGKAIICKNGISKRTFEIEKEIASCGVAKAIIINLAVYKQNR 62  
Db 32 FLLGLPLFGWFTKSNATNITTPGNMKAPLDELKAE-----NIKELYNF----- 77  
Qy 63 SYERLLALVDTVGRISGSK--NLEKAIQIMQNLQODGLEKYHLEPV-----RIPHW 113  
Db 78 -----TQIPLHAGTEQNFOIAKOIOSQWKEFGDLSVELAHYDVLLSYPNKTHPNY 127  
Qy 114 -----ERGEE--SAVMEPRRIHKIAIILGSSICTPPEGITAEVLYVTSFDELQRRASEA 166  
Db 128 ISIINEDENETFNISLFEF-----PPPEYENVSIVPPFSAFSPQGM-P 170  
Qy 167 RGIIVVYNQPIYINSRT-----VOYRTOGAWEAAKV-----ALASLIRSV 207  
Db 171 EGDLY-----YVNAKTEDPFKLERDMKINCSCGIIVARIGKVRGNKYNQAOLAGAKGV 225  
Qy 208 ASFS-----IYSPHTGIEYOD----- 224  
Db 226 ILYSDPADYFAP--GVKSYDPDGMNLPGGCVQGRNIIILNAGADPLTPGPANEYAYRRGI 283  
Qy 225 -----GVKPIPTACTIVEDAEMSRMASHGIKIYIQLKMGAKTYPD----- 265  
Db 284 AEAVGLPSIPVHPPIGYDDAKLLE-----KMGGSAPPDSSWRGSLKVPYNVGP 331  
Qy 266 -----TDSFNTVAETGSKYPEQVVLVSGHLDSDMDVGQAGMDG 304  
Db 332 GFTGNFSTQVKYKMHISTNEVTRIVNIGTLRGAVEPRRYVILGHRDSWVF--GGIDPQ 389  
Qy 305 GGAFISWE--ALSILKDLGLRPKRTLRVLVMTAEQGVGAFOYYQLHKVNISNYSLVW 361  
Db 390 SGAIVVHEIVRSFGLTKKEGWRPRTILFASWDAEEFGLGSTEWAE-----ENSRLLQ 443  
Qy 362 ESDAGTFLPTGLQTFGSEKARAMEEW-SLLQPLNITOVLSHGCTDINFWIQAIGVGA 420  
Db 444 ERGV-AYINADSSIEGNYTLRVDCPTPLMYSLVH--NLTKELKSPD-----EGFEGK 491  
Qy 421 SLDDLYK 428  
Db 492 SLYESWTK 499

RESULT 12  
US-09-78-295A-618  
; Sequence 618, Application US/09978295A  
; Patent No. US20020156006A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gottlieb, Mary E.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kijaviri, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James J.  
APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann  
 APPLICANT: Shelton, David L.  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Williams, P. Mickey  
 APPLICANT: Wood, William I.  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 FILE REFERENCE: P2630P1C11  
 CURRENT APPLICATION NUMBER: US/09/978,295A  
 CURRENT FILING DATE: 2001-10-15  
 PRIOR APPLICATION NUMBER: 09/918585  
 PRIOR FILING DATE: 2001-07-30  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/064249  
 PRIOR FILING DATE: 1997-11-03  
 PRIOR APPLICATION NUMBER: 60/065311  
 PRIOR FILING DATE: 1997-11-13  
 PRIOR APPLICATION NUMBER: 60/066364  
 PRIOR FILING DATE: 1997-11-21  
 PRIOR APPLICATION NUMBER: 60/077450  
 PRIOR FILING DATE: 1998-03-10  
 PRIOR APPLICATION NUMBER: 60/077632  
 PRIOR FILING DATE: 1998-03-11  
 PRIOR APPLICATION NUMBER: 60/077641  
 PRIOR FILING DATE: 1998-03-11  
 PRIOR APPLICATION NUMBER: 60/077649  
 PRIOR FILING DATE: 1998-03-11  
 PRIOR APPLICATION NUMBER: 60/077791  
 PRIOR FILING DATE: 1998-03-12  
 PRIOR APPLICATION NUMBER: 60/078004  
 PRIOR FILING DATE: 1998-03-13  
 PRIOR APPLICATION NUMBER: 60/078886  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/078936  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/078939  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/079294  
 PRIOR FILING DATE: 1998-03-25  
 PRIOR APPLICATION NUMBER: 60/079656  
 PRIOR FILING DATE: 1998-03-26  
 PRIOR APPLICATION NUMBER: 60/079664  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079689  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079663  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079728  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079786  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079920  
 PRIOR FILING DATE: 1998-03-30  
 PRIOR APPLICATION NUMBER: 60/079923  
 PRIOR FILING DATE: 1998-03-30  
 PRIOR APPLICATION NUMBER: 60/080105  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080107  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080165  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080194  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080327  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080328  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080333  
 PRIOR FILING DATE: 1998-04-01

PRIOR APPLICATION NUMBER: 60/080334  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/081070  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081049  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081071  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081195  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081955  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081819  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081952  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081838  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082568  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082569  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082704  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082804  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082700  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082797  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082796  
 PRIOR FILING DATE: 1998-04-23  
 PRIOR APPLICATION NUMBER: 60/083336  
 PRIOR FILING DATE: 1998-04-27  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083392  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083495  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083496  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083499  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083554  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083558  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083559  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083500  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083742  
 PRIOR FILING DATE: 1998-04-30  
 PRIOR APPLICATION NUMBER: 60/084366  
 PRIOR FILING DATE: 1998-05-05  
 PRIOR APPLICATION NUMBER: 60/084414  
 PRIOR FILING DATE: 1998-05-06  
 PRIOR APPLICATION NUMBER: 60/084441  
 PRIOR FILING DATE: 1998-05-06  
 PRIOR APPLICATION NUMBER: 60/084637  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084639  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084640

PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085689  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697

Query Match 6.7% Score 163; DB 10; Length 750;

Best Local Similarity 19.2%; Pred. No. 7, 9e-07; Indels 202; Gaps 23;  
Matches 105; Conservative 67; Mismatches 174;

QY 3 FLIFAFGVLHSLSCGKAIKNGISKRTPEEIKETIASCGDVAKAIINLAIVGKAQNR 62  
DB 32 FLIFGFLGWMFLIKSENENATNITPKNNMAFLDELAF-----NIKKFLHNF----- 77  
QY 63 SYERLALLVDYVGPLSGSK--NIEKAIQIMYQNLQODGLEKYLEV-----RIHHW 113  
DB 78 -----TQIPHLAGEQNFOQAKOIQSQWKEFGDSEVLAHYDVLISYNNKTHBNY 127  
QY 114 -----ERGBE--SAVMEPRRIKIAIILGSSIGPPEGITAEVLVYMSFDELQRASEA 166  
DB 128 ISIINEDGNEIFNLSLFE-----PPGEYENVSQIVPEFSAFSPQGM-P 170  
QY 167 RGIIVVYQPYINVSRT-----VOYRTQAVEAAKVG-----ALASLIRSV 207  
DB 171 BGDLY-----YVNVARTEDPFLERDMKINSGKIVIAHYGVFRGNKVKNAQLAGAKGV 225  
QY 208 ASFS-----TSPHTIGIYQD----- 224  
DB 226 ILYSPADYFAP--GVKSYPDGMNLPGGGVGRGNILINAGADPLTPGPAVEYARRGI 283  
QY 225 -----GVKIPRACTIVEDAEMMSRMAHSHGIXIVILKMGAKTYPD----- 265  
DB 284 AEAVGLPSIPVHPPIGYDAQKLE-----KKGSGAPPSWSRSGIKPYPNVGP 331  
QY 266 -----TDSFNTVAETGSKYPCQVVLVSGHLSMDVQGAMDDG 304  
DB 332 GFTGNFSTQKVMHISTNEVTRINVGTLGAVEPRVYVILGHRDSWVF--CGIDPQ 389  
QY 305 GGAFTSWE--ALSLIKDLGGLPKKTLRLVMTAEQGVGAFOYQLHKVNISVSLVM 361  
DB 390 SGAADVHIVSFGTLKKEGMPRTIIFASWDAEEFGLGSTEWAE-----ENSRLIQ 443  
QY 362 ESDAGTFLPTGQFGSGEKARAMEVW--SLQPLNITQVLSHGSGTINIMIQGVGA 420  
DB 444 ERGVAYINADSSIEGNTLRVDCTPLMYSLVH--NLTKELKSPD-----EGFEKG 491  
QY 421 SLDDLYK 428  
|| : |

DB 492 SLYESWTK 499  
RESULT 13  
US-09-978-697-618  
Sequence 618, Application US/09978697  
Patent No. US20020169284A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Deenoyers, Luc  
APPLICANT: Baton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C27  
CURRENT APPLICATION NUMBER: US/09/978,697  
CURRENT FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656



PRIOR FILING DATE: 1998-03-26  
 PRIOR APPLICATION NUMBER: 60/079664  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079689  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079663  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079728  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079786  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079920  
 PRIOR FILING DATE: 1998-03-30  
 PRIOR APPLICATION NUMBER: 60/079923  
 PRIOR FILING DATE: 1998-03-30  
 PRIOR APPLICATION NUMBER: 60/080105  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080107  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080165  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080194  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080327  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080328  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080333  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080334  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/081070  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081049  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081071  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081195  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081955  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081819  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081952  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081838  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082568  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082569  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082704  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082804  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082700  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082797  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082796  
 PRIOR FILING DATE: 1998-04-23  
 PRIOR APPLICATION NUMBER: 60/083336  
 PRIOR FILING DATE: 1998-04-27  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083392  
 PRIOR FILING DATE: 1998-04-29

PRIOR APPLICATION NUMBER: 60/083495  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083496  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083499  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083554  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083558  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083559  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083500  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083742  
 PRIOR FILING DATE: 1998-04-30  
 PRIOR APPLICATION NUMBER: 60/084366  
 PRIOR FILING DATE: 1998-05-05  
 PRIOR APPLICATION NUMBER: 60/084414  
 PRIOR FILING DATE: 1998-05-06  
 PRIOR APPLICATION NUMBER: 60/084441  
 PRIOR FILING DATE: 1998-05-06  
 PRIOR APPLICATION NUMBER: 60/084637  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084639  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084640  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084598  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084600  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084627  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084643  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/085339  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085338  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085323  
 PRIOR FILING DATE: 1998-05-13  
 PRIOR APPLICATION NUMBER: 60/085582  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085700  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085689  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085579  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085580  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085573  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085704  
 PRIOR FILING DATE: 1998-05-15  
 PRIOR APPLICATION NUMBER: 60/085697

Query Match 6.7%; Score 163; DB 10; Length 750;  
 Best Local Similarity 19.2%; Pred. No. 7.9e-07;  
 Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

QY 3 FLTAFPGVHLISLCSGKAI CKNGISKRTFEIEKIEIASCQDVAKAIIMLAVYGAQNR 62  
 DB 32 FPLGLFPGWPKISSNATNITPRHNKKAFLELKA-----NKKFLHNF----- 77  
 QY 63 SYERLALVDTGGRSLGSK---NLEKAIQIMTONLOODGLEKXHEPV-----RIPHW 113  
 DB 78 -----TQIPHLAGTEONFOLAKOIOSOWMEFGLDVEVLAHYVLLSYPKKTHPNY 127  
 QY 114 -----ERGEE--SAVMLEPRTHKIALIIGLSSISCTPREGITAEVLVVTSPDELQRASEA 166

Db 128 ISIINEDGNEIFNTSLFEP-----PPGXYENVSDIVPPFAFSPQW-P 170  
Qy 167 RGIYVNVQPIYINERT-----VOYRTOGAVEAAKV-----ALASLISV 207  
Db 171 EGDIV-----YVNVARTDEFFLEERDMKINCSGKIYIARYGVFRGNKYKNAQLGAGV 225  
Qy 208 ASFS-----IYSPHTGIXOYD----- 224  
Db 226 ILVSPDAPYFAP--GVKSTYPOGMNIPGGGVORGNILINAGADPLITPGYPAINEYVRGI 283  
Qy 225 ---GVKPIPTACTIVEDAEMMSRMAHGIKIVIQKNGAKTYPD----- 265  
Db 284 AEAVALPSIPVHPGIVYDAQKLB-----KKGSGAPDSSWRGSLKVPYVNGP 331  
Qy 266 -----TDSNTVAETGSKYPEQVYLVSHLHDSMDVGQAMDG 304  
Db 332 GFTGNFSTQKVMKIHSTNEVTRINVIQTLLGAVEPRRYVLLGHRDSWPF--GIDPQ 389  
Qy 305 GGAFTSWE--ALSLIKDLGLRPKRTLRLVLTAEQGVGAFOYYOLHKVNISYSLVM 361  
Db 390 SGAAVNHVIVRSGFLKKEGMRPRRTILFASMDAEFFGLISTEMAF-----EKSRLIQ 443  
Qy 362 ESDACTFPTGLOFTGSEKARAMEVM-SLLOPLNITQVLSHGEGTDINFWIOAGVPGA 420  
Db 444 ERGV-AVINADSSIGNVTLRVDCTPLMYSLVH--NLTFELKSPD-----EGFEKG 491  
Qy 421 SLUDDLYK 428  
Db 492 SLYESWTK 499

## RESULT 14

US-09-978-192A-618

Sequence 618, Application US/09978192A

Patent No. US20020177553A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi

APPLICANT: Baker Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan

APPLICANT: Ferrara, Napoleon

APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Grimaldi, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J.

APPLICANT: Kijavlin, Ivar J.

APPLICANT: Kuo, Sophia S.

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Shelton, David L.

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

FILE REFERENCE: P263091C9

CURRENT APPLICATION NUMBER: US/09/978,192A

PRIOR FILING DATE: 2001-10-15

PRIOR FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

;; PRIOR APPLICATION NUMBER: 60/064249  
;; PRIOR FILING DATE: 1997-11-03  
;; PRIOR APPLICATION NUMBER: 60/065311  
;; PRIOR FILING DATE: 1997-11-13  
;; PRIOR APPLICATION NUMBER: 60/066364  
;; PRIOR FILING DATE: 1997-11-21  
;; PRIOR APPLICATION NUMBER: 60/077450  
;; PRIOR FILING DATE: 1998-03-10  
;; PRIOR APPLICATION NUMBER: 60/077632  
;; PRIOR FILING DATE: 1998-03-11  
;; PRIOR APPLICATION NUMBER: 60/077641  
;; PRIOR FILING DATE: 1998-03-11  
;; PRIOR APPLICATION NUMBER: 60/077649  
;; PRIOR FILING DATE: 1998-03-11  
;; PRIOR APPLICATION NUMBER: 60/077791  
;; PRIOR FILING DATE: 1998-03-12  
;; PRIOR APPLICATION NUMBER: 60/078004  
;; PRIOR FILING DATE: 1998-03-13  
;; PRIOR APPLICATION NUMBER: 60/078886  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078936  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078939  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/079656  
;; PRIOR FILING DATE: 1998-03-26  
;; PRIOR APPLICATION NUMBER: 60/079664  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079689  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079663  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079728  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079786  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079920  
;; PRIOR FILING DATE: 1998-03-30  
;; PRIOR APPLICATION NUMBER: 60/079923  
;; PRIOR FILING DATE: 1998-03-30  
;; PRIOR APPLICATION NUMBER: 60/080105  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080107  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080165  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080194  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080327  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080328  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080333  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080334  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/081070  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081049  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081071  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081195  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081955

[illegible]

APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James;  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P26301C63  
CURRENT APPLICATION NUMBER: US/09/999,832A  
CURRENT FILING DATE: 2001-10-24  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30

PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500

PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085689  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697

Query Match 6.7%; Score 163; DB 10; Length 750;  
Best Local Similarity 19.2%; Pred. No. 7.9e-07;  
Matches 105; Conservative 67; Mismatches 174; Indels 202; Gaps 23;

QY 3 FLIFAFGCVHLSICSGKAIKNGISKRTFBEIKKEIASCGDVAKAIINLAVYGAQNR 62  
DB 32 FFLIGFLFGWFIKSNENATNITPKANMKAFLDELKAE-----NIKKFLHNF----- 77  
QY 63 SYERIALLVDTVGPRLSGSK--NLEKAIQIMYQNLQODGLEKXVLEPV-----RIPHW 113  
DB 78 -----TQIPHLAQTGQNFQAKQIQSQWKEFGDSVELAHYDVLISYPNKTHTPNY 127  
QY 114 -----ERGEF--SAVMEPRRIHKIAIILGSSIGTPPEGCITAEVLVVTSPDELQRRASEA 166  
DB 128 ISIINEDNEIEFNISLFEF-----PPGYENVSIDIVPPFSAFSPQGM-P 170  
QY 167 RGIIVVYVQPIYNSRT-----VOYRTQGAVEAAKVG-----ALASLIRSV 207  
DB 171 EGDLV-----YVNAKRTDFPKLERDMKINCSGKIIVIRYGVFRGNKVNKAQJLAGAKGV 225  
QY 208 ASFS-----IYSPHTGIGOEYOD----- 224  
DB 226 ILVSDPADYFAP--GVKSYPDGMNLPGGVQGRGNILNLNGAGDPLTPGYPANEVAYRRGI 283  
QY 225 -----GVKPIPTNACTIVEDAEEMMSKMAHSGIKIVIQIKMGAKTYPD----- 265  
DB 284 AEAVALPSIPIVPIGIDYDAOKLLE-----KMGSAAPDSSWRGSLKVPYVNVGP 331

QY 266 -----TDSFNTVAEITGSKYPEQVVLVSGHLSNDVQGAAMDG 304  
DB 332 GFTGNFSTQKVKMHISTNEVTIRIYVIGTLRGAVEPDRYVILGHRDSWVF--GGIDPQ 389  
QY 305 GGAFTSWE--ALSIXDGLRPKRTLRVLMTAEBOGGVGAQOYQJLHKVNISNLSVM 361  
DB 390 SGAIVHEIVRSFGLTKKEGMRPRRTILFASMDAEFEGLIGSTEMAE-----ENSRLLQ 443  
QY 362 ESDAGTFLPTGLQFTGSEKARAIMEEVM--SLQPLNITOVLSHGEGTDINFWIQAGVPGA 420  
DB 444 ERGV-AYINADSSIEGNYTLRVDCPTPLMYSLVH--NLTKEKSPD-----EGFEKG 491  
QY 421 SLDDLYK 428  
DB 492 SLYESWTK 499

Search completed: December 22, 2003, 11:54:05  
Job time: 42.3941 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:34:19 ; Search time 29.7527 Seconds  
(without alignments)  
933.601 Million cell updates/sec

Title: US-09-745-763-36\_COPY\_1\_175

Perfect score: 883

Sequence: 1 MKELIFAFGCVHLSTSCG.....FDELQRRASPARKIVVYQ 175

Scoring table: BLAST62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	883	100.0	472	18	AAW33604
2	883	100.0	472	19	AAW85456
3	883	100.0	472	21	AAAB10229
4	883	100.0	472	21	AAAB7254
5	883	100.0	472	21	AAAB879
6	883	100.0	472	22	AAAB6058
7	883	100.0	472	23	ABP61800
8	883	100.0	472	23	ABG33880
9	883	100.0	472	24	ABR47892

10	883	100.0	472	24	ABR48106
11	883	100.0	472	24	ABR00152
12	883	100.0	472	24	ABR00274
13	883	100.0	473	21	AAAB7081
14	883	100.0	473	21	ABG09405
15	883	100.0	472	22	AAAB25762
16	883	100.0	472	22	AAAB93559
17	460	52.1	149	18	AAW27642
18	460	52.1	149	18	AAW31633
19	460	52.1	149	18	AAW44078
20	397	45.0	90	22	AAAB4050
21	351	39.8	71	20	AAV11865
22	111	12.6	33	20	AAV11861
23	89	10.1	34	22	ABG09404
24	85.5	9.7	37	22	AAAB82916
25	82	9.3	524	22	AAW40332
26	82	9.3	533	21	AAAB43002
27	82	9.3	533	21	AAW40333
28	81.5	9.2	162	19	AAW98864
29	81	9.2	340	24	ABP56996
30	81	9.2	809	23	ABB90610
31	81	9.2	812	20	AAV35560
32	80.5	9.1	222	22	AAU31028
33	79.5	9.0	406	23	ABP38639
34	79.5	9.0	1863	16	AAV76641
35	79	8.9	1216	21	AAAB42097
36	79	8.9	1390	23	ABG96292
37	78.5	8.9	288	17	AAAB1439
38	78.5	8.9	308	21	AAAB1246
39	78.5	8.9	312	17	AAAB99442
40	78.5	8.9	312	18	AAW10005
41	78.5	8.9	312	19	AAW79891
42	78.5	8.9	315	21	AAAB1245
43	78.5	8.9	344	17	AAAB1501
44	78.5	8.9	351	21	AAAB1244
45	78.5	8.9	370	23	AAO22591

# ALIGNMENTS

RESULT 1	AAW33604	AAW33604 standard; Protein; 472 AA.
AC	AAW33604;	
XX		
DT	21-MAY-1998 (first entry)	
XX		
DE	Human secreted protein AM282 full-length sequence.	
XX		
KW	Secreted protein; AM282; cytokine; human.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	Peptide	1..24
FT	Protein	/label= sig_peptide
XX		25..472
XX		/label= Mat_protein
XX		
XX	WO9739030-A2.	
XX	23-OCT-1997.	
XX		
XX	16-APR-1997;	97WO-US06475.
XX		
XX	13-JAN-1997;	97US-0783520.
XX	18-APR-1996;	96US-0634325.
XX		
XX	(GENY ) GENETICS INST INC.	
XX		
XX	Jacobs K, LaVallie ER, McCoy JM, Werberg D, Racie LA;	

PI Spaulding V;  
 XX  
 DR WPI: 1997-526400/48.  
 DR N-PSDB; AAV02296.  
 XX  
 FT New isolated secretory proteins AM340, AM282 and AK583 - possibly  
 PT have cytokine, cell proliferation/differentiation regulating,  
 PT immunomodulating activities, etc.  
 XX  
 PS Claim 16; Page 45-47; 59pp; English.  
 XX  
 CC This human secreted protein, designated AM282, is encoded by a  
 CC full-length cDNA clone (see AAV02296), deposited in ATCC 98026, that  
 CC was identified from a database search using an isolated partial  
 CC AM282 clone (see AAT97398). AM282 protein can be used in a claimed  
 CC method for preventing, treating or ameliorating a medical  
 CC condition. It may exhibit cytokine, cell proliferation (either  
 CC inducing or inhibiting) or cell differentiation (either inducing or  
 CC inhibiting) activity or may induce production of other cytokines in  
 CC certain cell populations. It may also exhibit e.g. immune  
 CC stimulating or suppressing activity, haematopoiesis regulating  
 CC activity, tissue growth activity, activin/inhibin activity,  
 CC chemotactic or chemokinetic activity, haemostatic or thrombolytic  
 CC activity, receptor/ligand activity, anti-inflammatory activity,  
 CC tumour inhibition activity, or other activities. No evidence of  
 CC any of these activities is given in the specification.  
 XX  
 SQ Sequence 472 AA;  
 Query Match 100.0%; Score 883; DB 18; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 2,2e-83;  
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKPLIFAFPGVHLLSLCSGKAICNGISKRTPEEIKKEIASCGVAAKAIINLAVYKQAQ 60  
 DB 1 MKPLIFAFPGVHLLSLCSGKAICNGISKRTPEEIKKEIASCGVAAKAIINLAVYKQAQ 60  
 QY 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYONLQODGLEKVLPEVRIPIHMERGESA 120  
 DB 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYONLQODGLEKVLPEVRIPIHMERGESA 120  
 QY 121 VMLERPIHKIAIILGSSIGTPPEGITAEVLVVTSPFDELORRASEARGKIVVYNO 175  
 DB 121 VMLERPIHKIAIILGSSIGTPPEGITAEVLVVTSPFDELORRASEARGKIVVYNO 175

RESULT 2  
 AAW85456  
 ID AAW85456 standard; Protein; 472 AA.  
 XX  
 AC AAW85456;  
 XX  
 DT 25-FEB-1999 (first entry)  
 XX  
 DE Secreted protein encoded by clone bu45\_2.  
 XX  
 KW Secreted protein; nutritional activity; immune stimulating; vaccine;  
 KW suppressing activity; haematopoiesis regulating activity;  
 KW tissue growth activity; activin; inhibin activity; chemotactaxis;  
 KW chemokinetic activity; haemostasis; thrombolytic activity; receptor;  
 KW ligand; anti-inflammatory; cadherin; tumour invasion suppressor;  
 KW tumour inhibition; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9842739-A2.  
 XX  
 PD 01-OCT-1998.  
 XX  
 PF 20-MAR-1998; 98WO-US05653.  
 XX  
 PR 19-MAR-1998; 98US-0044466.  
 PR 21-MAR-1997; 97US-0822167.

XX  
 PA (GENY ) GENETICS INST INC.  
 XX  
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;  
 PI Racie LA, Spaulding V, Treacy M;  
 XX  
 DR WPI: 1998-609890/51.  
 DR N-PSDB; AAV82779.  
 XX  
 PT New polynucleotides encoding secreted human proteins - derived from  
 PT human foetal brain, adult brain, foetal kidney, placenta or adult  
 PT pineal gland cDNA libraries.  
 XX  
 PS Claim 14; Page 70-72; 113pp; English.  
 XX  
 CC The present sequence represents a secreted protein. The polynucleotide  
 CC and secreted protein are predicted to have biological activities which  
 CC would make them suitable for treating, preventing or ameliorating medical  
 CC conditions in humans and animals, although no supporting data is given.  
 CC Suggested activities include nutritional activity, immune stimulating  
 CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating  
 CC activity, tissue growth activity, activin/inhibin activity,  
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,  
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour  
 CC invasion suppressor activity, and tumour inhibition activity (no data is  
 CC given in the specification to support these activities). The  
 CC polynucleotide is also stated to be useful for gene therapy.  
 XX  
 SQ Sequence 472 AA;  
 Query Match 100.0%; Score 883; DB 19; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 2,2e-83;  
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKPLIFAFPGVHLLSLCSGKAICNGISKRTPEEIKKEIASCGVAAKAIINLAVYKQAQ 60  
 DB 1 MKPLIFAFPGVHLLSLCSGKAICNGISKRTPEEIKKEIASCGVAAKAIINLAVYKQAQ 60  
 QY 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYONLQODGLEKVLPEVRIPIHMERGESA 120  
 DB 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYONLQODGLEKVLPEVRIPIHMERGESA 120  
 QY 121 VMLERPIHKIAIILGSSIGTPPEGITAEVLVVTSPFDELORRASEARGKIVVYNO 175  
 DB 121 VMLERPIHKIAIILGSSIGTPPEGITAEVLVVTSPFDELORRASEARGKIVVYNO 175

RESULT 3  
 AAB10229  
 ID AAB10229 standard; Protein; 472 AA.  
 XX  
 AC AAB10229;  
 XX  
 DT 16-NOV-2000 (first entry)  
 XX  
 DE Human fetal kidney protein fragment AM282\_11.  
 XX  
 KW Secreted protein; cytostatic; immunostimulatory; antimicrobial;  
 KW antiviral; immunosuppressive; anti-inflammatory; vulnery; cytokine;  
 KW cell proliferation; differentiation; regulator; treatment; tumor;  
 KW autoimmune disease; inflammatory disorder; wound; microbial infection;  
 KW viral disease; graft versus host reaction suppression.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200037630-A1.  
 XX  
 PD 29-JUN-2000.  
 XX  
 PF 22-DEC-1999; 99WO-US31005.  
 XX  
 PR 23-DEC-1998; 98US-0220876.



PA (GENY) GENETICS INST INC.  
 XX  
 PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Bowman MR;  
 XX  
 DR WPI; 2000-442661/38.  
 DR N-PSDB; AAA40493.  
 XX  
 PT Secreted human proteins AS296-1i and AS34-1i, useful for treating  
 PT tumors, autoimmune diseases, inflammatory disorders, wounds, microbial  
 PT infections and viral diseases -  
 XX  
 PS Disclosure; Page 199-200; 293pp; English.  
 XX  
 CC This invention describes novel secreted human proteins (I) which have  
 CC cytostatic, immunostimulatory, antimicrobial, antiviral,  
 CC immunosuppressive, anti-inflammatory and vulnereary activity and which act  
 CC as cytokine, cell proliferation or differentiation regulators. (I)  
 CC is useful for treating tumors, autoimmune diseases, inflammatory  
 CC disorders, wounds, microbial infections and viral diseases. (I) is also  
 CC useful for suppressing graft versus host reaction. AAB10226-B10288  
 CC represent the secreted proteins encoded by AAA40490-A40580 which are  
 CC described in the method of the invention.  
 CC  
 XX  
 SQ Sequence 472 AA;  
 Query Match 100.0%; Score 883; DB 21; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-83;  
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKFLIFAFPGVHLISCSGKAIKNGISKRTPEEIKKEIASCSDVAKAIINLAIVGKAQ 60  
 DB 1 MKFLIFAFPGVHLISCSGKAIKNGISKRTPEEIKKEIASCSDVAKAIINLAIVGKAQ 60  
 QY 61 NRSYERIALLVDTYVGPRLSSKNLEKAIQIMYQNLQDGLKYLEPVRIPHERGESBA 120  
 DB 61 NRSYERIALLVDTYVGPRLSSKNLEKAIQIMYQNLQDGLKYLEPVRIPHERGESBA 120  
 QY 121 VMLEPRIHKIALIIGLSSIGTPPEGITAEVAVVTSFDELORRASEARGKIVVNO 175  
 DB 121 VMLEPRIHKIALIIGLSSIGTPPEGITAEVAVVTSFDELORRASEARGKIVVNO 175  
 DE Human signal peptide containing protein HSP-31 SEQ ID NO:31.  
 XX  
 KW Human; signal peptide-containing protein; HSP; diagnosis; cancer;  
 KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;  
 KW antimicrobial; neuroprotective; cardiovascular; hepatotropic;  
 KW antiaesthetic; gene therapy; cell proliferation; neurological disorder;  
 KW reproductive disorder; developmental disorder; arteriosclerosis;  
 KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;  
 KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;  
 KW Parkinson's disease; Huntington's diseases; ovulatory defect;  
 KW muscular dystrophy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200000610-A2.  
 XX  
 PD 06-JAN-2000.  
 XX  
 PF 25-JUN-1999; 99WO-US14484.  
 XX  
 PR 26-JUN-1998; 98US-0090762.  
 PR 31-JUL-1998; 98US-0094983.

PR 01-OCT-1998; 98US-0102686.  
 PR 11-DEC-1998; 98US-0112129.  
 XX  
 PA (INCYTE) INCYTE PHARM INC.  
 XX  
 PI Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;  
 PI Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;  
 PI Bandman O;  
 XX  
 DR WPI; 2000-160673/14.  
 DR N-PSDB; AA298139.  
 XX  
 PT New human signal peptide-containing proteins useful in treatment,  
 PT prevention and diagnosis of e.g. cancer, inflammation and  
 PT cardiovascular disease -  
 XX  
 PS Claim 1; Page 179-180; 327pp; English.  
 XX  
 CC AA298109 to AA298242 encode AAY87224 to AAY87357 which represent the  
 CC human signal peptide-containing proteins HSP-1 to HSP-134. HSPs have  
 CC anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatotropic,  
 CC neuroprotective, cardiovascular and antiaesthetic activities, and can  
 CC be used in gene therapy. HSPs can be used to treat or prevent disorders  
 CC associated with decreased activity or function of HSP. Antagonists of  
 CC HSP are used to treat or prevent disorders associated with increased  
 CC activity or function of HSP. Such disorders include cell proliferation  
 CC (including cancer), inflammation, cardiovascular, neurological,  
 CC reproductive or developmental disorders, (e.g. arteriosclerosis,  
 CC cirrhosis, psoriasis, acquired immune deficiency syndrome, congestive or  
 CC ischaemic heart disease, microbial or other infections, anaemia or  
 CC diseases, schizophrenia, Alzheimer's, Parkinson's or Huntington's  
 CC disease, muscular dystrophy). HSP  
 CC nucleic acids can be used for the recombinant production of HSP, for  
 CC detecting HSP in standard hybridisation and amplification assays (for  
 CC diagnosis and monitoring), in gene therapy, as antisense,  
 CC triplex-forming or ribozyme therapeutics, for detecting related sequences  
 CC or genetic variations, and for chromosomal mapping. HSP are also used to  
 CC raise specific antibodies (Ab) and to screen for agonists and  
 CC antagonists (potential therapeutic agents). Ab are used to diagnose, or  
 CC monitor, HSP-related diseases (in usual immunoassays), as therapeutic  
 CC antagonists, in competitive drug screens, and for purification of HSP  
 CC from natural sources.  
 CC  
 XX  
 SQ Sequence 472 AA;  
 Query Match 100.0%; Score 883; DB 21; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-83;  
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKFLIFAFPGVHLISCSGKAIKNGISKRTPEEIKKEIASCSDVAKAIINLAIVGKAQ 60  
 DB 1 MKFLIFAFPGVHLISCSGKAIKNGISKRTPEEIKKEIASCSDVAKAIINLAIVGKAQ 60  
 QY 61 NRSYERIALLVDTYVGPRLSSKNLEKAIQIMYQNLQDGLKYLEPVRIPHERGESBA 120  
 DB 61 NRSYERIALLVDTYVGPRLSSKNLEKAIQIMYQNLQDGLKYLEPVRIPHERGESBA 120  
 QY 121 VMLEPRIHKIALIIGLSSIGTPPEGITAEVAVVTSFDELORRASEARGKIVVNO 175  
 DB 121 VMLEPRIHKIALIIGLSSIGTPPEGITAEVAVVTSFDELORRASEARGKIVVNO 175  
 DE Human signal peptide containing protein HSP-31 SEQ ID NO:31.  
 XX  
 KW Human; signal peptide-containing protein; HSP; diagnosis; cancer;  
 KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;  
 KW antimicrobial; neuroprotective; cardiovascular; hepatotropic;  
 KW antiaesthetic; gene therapy; cell proliferation; neurological disorder;  
 KW reproductive disorder; developmental disorder; arteriosclerosis;  
 KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;  
 KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;  
 KW Parkinson's disease; Huntington's diseases; ovulatory defect;  
 KW muscular dystrophy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200000610-A2.  
 XX  
 PD 06-JAN-2000.  
 XX  
 PF 25-JUN-1999; 99WO-US14484.  
 XX  
 PR 26-JUN-1998; 98US-0090762.  
 PR 31-JUL-1998; 98US-0094983.

KW Chromosome 8q21.3; prostate cancer; neurodegenerative disease;  
 KW Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;  
 KW peripheral neuropathy; Huntington's disease; acute brain injury;  
 KW multiple sclerosis; peripheral nerve trauma; ischaemia; dementia;  
 KW therapy; diagnosis; neurotrophic; neuroprotective; neuroleptic;  
 KW antiparkinsonian; anticonvulsant; vasodilator.  
 OS Homo sapiens.  
 XX  
 XX  
 XX Key Location/Qualifiers  
 FH Domain 3..24  
 FT /note= "membrane-spanning domain"  
 FT Modified-site 121  
 FT /note= "N-glycosylated"  
 FT Modified-site 179  
 FT /note= "N-glycosylated"  
 FT Modified-site 353  
 FT /note= "N-glycosylated"  
 FT Modified-site 356  
 FT /note= "N-glycosylated"  
 FT Modified-site 396  
 FT /note= "N-glycosylated"  
 FT  
 FT  
 FN WO200004157-A2.  
 XX  
 XX 27-JAN-2000.  
 XX  
 XX 14-JUL-1999; 99WO-GB02241.  
 XX  
 XX 14-JUL-1998; 98GB-0015284.  
 PR  
 PA (JANNC ) JANSSEN PHARM NV.  
 XX  
 XX Pangalos M, Neefs JEFM, Peeters DCG.  
 PI  
 DR WPI; 2000-182424/16.  
 DR N-PSDB; AAZ58313.  
 XX  
 PT New human N-acetylated alpha-linked acidic dipeptidases for treating  
 PT neural disorders e.g. Alzheimer's disease, schizophrenia and  
 PT Parkinson's disease -  
 XX  
 XX Claim 15; Fig 5; 95pp; English.  
 PS  
 XX The present sequence is that predicted for human N-acetylated  
 CC alpha-linked acidic dipeptidase IV (NALAD-ase IV) on the basis of  
 CC isolated cDNA (see AAZ58313). NALAD-ase IV is predicted to be a  
 CC type II integral membrane protein of mol.wt. 51.9 kDa and pI 5.99.  
 CC Expression was low in all tissues examined by RT-PCR. The  
 CC invention provides human NALAD-ase I, II and IV cDNAs and encoded  
 CC polypeptides, as well as vectors, host cells, transgenic organisms,  
 CC antisense nucleic acids, agonists and antagonists. These are useful  
 CC for treating neural disorders such as Alzheimer's disease,  
 CC schizophrenia, ALS, Parkinson's disease, peripheral neuropathy,  
 CC Huntington's disease, acute brain injury, multiple sclerosis,  
 CC exposure to neurotoxins, peripheral nerve trauma, ischaemia or  
 CC dementia (claimed).  
 CC  
 CC  
 SQ Sequence 472 AA;  
 Query Match 100.0%; Score 883; DB 21; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-83;  
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKFLIPAFGGVHLISLCSGKAICNGISKRTFEIEKEIASCGDVAKIINLAVYGAQ 60  
 DB 1 MKFLIPAFGGVHLISLCSGKAICNGISKRTFEIEKEIASCGDVAKIINLAVYGAQ 60  
 QY 1 NRSYERLALVDYGPRLSSGSKNLEKATQIWNQIQDGLRKVHLEPRIMHMERGESA 120  
 DB 1 NRSYERLALVDYGPRLSSGSKNLEKATQIWNQIQDGLRKVHLEPRIMHMERGESA 120  
 QY 61 NRSYERLALVDYGPRLSSGSKNLEKATQIWNQIQDGLRKVHLEPRIMHMERGESA 120  
 DB 61 NRSYERLALVDYGPRLSSGSKNLEKATQIWNQIQDGLRKVHLEPRIMHMERGESA 120  
 QY 121 VMEPRIHKAIILGSSISGTPPEGITAEVLVWTSFDELQRRASFGAKIVVYVQ 175

DB 121 VMEPRIHKAIILGSSISGTPPEGITAEVLVWTSFDELQRRASFGAKIVVYVQ 175  
 RESULT 6  
 AA06058  
 ID AA06058 standard; Protein; 472 AA.  
 XX  
 XX  
 AC AA06058;  
 XX  
 DT 24-SEP-2001 (first entry)  
 XX  
 DE Human gene 18 encoded secreted protein HRA03J5, SEQ ID NO:120.  
 XX  
 KW Human; secreted protein; proliferative disorder; cancer; tumour; asthma;  
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;  
 KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;  
 KW inflammation; neurological disorder; Alzheimer's disease; food additive;  
 KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;  
 KW pregnancy-related disorder; endocrine disorder; infection; wound healing;  
 KW cell culture; chemotaxis; vulnarity; binding partner identification;  
 KW gene therapy.  
 OS Homo sapiens.  
 XX  
 XX  
 XX Key Location/Qualifiers  
 FH Peptide 1..24  
 FT /label= signal\_peptide  
 FT Protein 25..472  
 FT /label= Mature\_human\_secreted\_protein  
 XX  
 XX WO200151504-A1.  
 XX  
 XX 19-JUL-2001.  
 XX  
 XX 12-JAN-2001; 2001WO-US00911.  
 XX  
 XX 13-JAN-2000; 2000US-0482273.  
 PR  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y;  
 PI Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;  
 PI Endress GA, Muscenski M, Emner R;  
 XX  
 XX WPI; 2001-425865/45.  
 DR N-PSDB; AAD11647.  
 XX  
 PT Isolated nucleic acid molecule encoding a human secreted protein is  
 PT used in preventing, treating or ameliorating a medical condition -  
 XX  
 XX Claim 11; Page 750-752; 864pp; English.  
 PS  
 XX AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted  
 CC protein genes, and AAD06041-AAD06132 represent the proteins they encode.  
 CC AAD06133-AAD06205 represent human secreted protein fragments.  
 CC The secreted proteins and their genes are useful for preventing, treating  
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
 CC Pathological conditions can be diagnosed by determining the amount of the  
 CC new protein in a sample or by determining the presence of mutations in  
 CC the new genes. Specific uses are described for each of the 71 genes,  
 CC based on the tissues in which they are most highly expressed, and include  
 CC developing products for the diagnosis or treatment of proliferative  
 CC disorders, cancer, tumours, foetal and developmental abnormalities,  
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,  
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,  
 CC pregnancy-related disorders, endocrine disorders, and infections. The

CC proteins can also be used to aid wound healing and epithelial cell  
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
 CC before transplantation, for supporting cell culture of primary tissues,  
 CC to regenerate tissues, to identify their cognate ligands or binding  
 CC partners, and in chemotaxis, and can be used as a food additive or  
 CC preservative to modify storage properties. Antibodies specific for a  
 CC protein of the invention can be used in alleviating symptoms associated  
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).  
 CC The present sequence represents a human secreted protein of  
 CC the invention.

XX Sequence 472 AA;

Query Match 100.0%; Score 883; DB 22; Length 472;

Best Local Similarity 100.0%; Pred. No. 2,2e-83; Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAFPGVHLLSLCSGKAICKNGISKRTPEEIKKEIASCGDVAKAIINLAVYGAQ 60

DB 1 MKFLIFAFPGVHLLSLCSGKAICKNGISKRTPEEIKKEIASCGDVAKAIINLAVYGAQ 60

QY 61 NRSYERLALLVDVTGPRLSGSKNLEKAIQIMYQNLQDGLKHLFVRIPHERGESA 120

DB 61 NRSYERLALLVDVTGPRLSGSKNLEKAIQIMYQNLQDGLKHLFVRIPHERGESA 120

QY 121 VMLFPRTHKTAIIGLSSIGTPPEGITAEVLVVTSPFELORRASEANGKIIVTNO 175

DB 121 VMLFPRTHKTAIIGLSSIGTPPEGITAEVLVVTSPFELORRASEANGKIIVTNO 175

RESULT 7

ABP61800 ID ABP61800 standard; Protein; 472 AA.

AC ABP61800;

DT 04-OCT-2002 (first entry)

DE Human polypeptide SEQ ID NO 154.

XX Human; cytostatic; antirheumatic; antiarthritic; vulnery; analgesic;  
 XX antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;  
 XX neuroprotective; nootropic; osteoprotic; haemostatic; vasotrophic;  
 XX antitumor; fungicide; antidiabetic; antiaslthmatic; antiallergic;  
 XX immunostimulant; antiparasitic; secreted protein; transmembrane protein;  
 XX cytokine; cell proliferation; cell differentiation; autoimmune disease;  
 XX stem cell; growth factor; nervous system disease; neuropathy;  
 XX Alzheimer's disease; Parkinson's disease; Huntington's disease;  
 XX osteoporosis; severe combined immunodeficiency; SCID; infection;  
 XX multiple sclerosis; rheumatoid arthritis; gene therapy.

OS Homo sapiens.

PN US2002065394-A1.

PD 30-MAY-2002.

PF 22-DEC-2000; 2000US-0745763.

PR 18-MAR-1998; 98US-0040963.

PA (JACO/) JACOBS K.

PA (MCCO/) MCCOY J M.

PA (LAVA/) LAVALLIE E R.

PA (COLL/) COLLINS-RACIE L A.

PA (EVAN/) EVANS C.

PA (MERB/) MERBERG D.

PA (TREA/) TREACY M.

PA (SPAUV/) SPAULDING V.

PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C, Merberg D, Treacy M, Spaulding V;

XX WPI; 2002-582343/62.

DR N-PSDB; ABQ92016.

PT Novel secreted or transmembrane protein and polynucleotide encoding the

PT protein, useful for diagnosis and treatment of neurological disorders,

PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis

PS Claim 51; Page 114-115; 284pp; English.

CC The invention relates to human secreted or transmembrane protein (I),  
 CC their fragments and is encoded by specific complementary deoxyribonucleic  
 CC acid (CDNA) inserts (II), where the protein is substantially free from  
 CC other mammalian proteins. (I) are useful for preventing, treating or  
 CC ameliorating a medical condition, especially immunological treatment or  
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,  
 CC cytokine, cell proliferation, cell differentiation, antiinflammatory,  
 CC stem cell growth factor activity and activin or inhibin-related  
 CC activities. (I) can be used to manipulate stem cells in culture to give  
 CC rise to neuroepithelial cells that can be used to augment or replace  
 CC cells damaged by illness, autoimmune disease, accidental damage or  
 CC genetic disorders. (I) induces the proliferation of neural cells and  
 CC regeneration of nerve and brain tissue and is useful for the treatment of  
 CC central and peripheral nervous system diseases and neuropathies, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic  
 CC activity, regulation of haematopoiesis and is useful for treating myeloid  
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopaenia  
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,  
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or  
 CC periodontal disease. (I) is also useful for gut protection or  
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury  
 CC in various tissues, various immune deficiencies and disorders including  
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,  
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,  
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,  
 CC such as asthma or other respiratory problems. (II) is useful to express  
 CC recombinant protein, as markers for tissues in which the corresponding  
 CC protein is preferentially expressed and in gene therapy. The present  
 CC sequence is that of a polypeptide of the invention.

SO Sequence 472 AA;

Query Match 100.0%; Score 883; DB 23; Length 472;

Best Local Similarity 100.0%; Pred. No. 2,2e-83; Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAFPGVHLLSLCSGKAICKNGISKRTPEEIKKEIASCGDVAKAIINLAVYGAQ 60

DB 1 MKFLIFAFPGVHLLSLCSGKAICKNGISKRTPEEIKKEIASCGDVAKAIINLAVYGAQ 60

QY 61 NRSYERLALLVDVTGPRLSGSKNLEKAIQIMYQNLQDGLKHLFVRIPHERGESA 120

DB 61 NRSYERLALLVDVTGPRLSGSKNLEKAIQIMYQNLQDGLKHLFVRIPHERGESA 120

QY 121 VMLFPRTHKTAIIGLSSIGTPPEGITAEVLVVTSPFELORRASEANGKIIVTNO 175

DB 121 VMLFPRTHKTAIIGLSSIGTPPEGITAEVLVVTSPFELORRASEANGKIIVTNO 175

RESULT 8

ABG33880 ID ABG33880 standard; Protein; 472 AA.

AC ABG33880;

DT 15-JUL-2002 (first entry)

DE Human secreted protein encoded by gene 18 #1.

XX Human; secreted protein; gene therapy; immunosuppressive;

KM antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;  
KM vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;  
KM virocidic; fungicide; ophthalmological; autoimmune disease; neoplasm;  
KM rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;  
KM cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;  
KM angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
KM ocular disorder; corneal infection; wound healing; skin aging;  
KM epithelial cell proliferation; food additive.  
XX  
XX Homo sapiens.  
XX WO200226931-A2.  
XX  
XX 04-APR-2002.  
XX  
XX 24-SEP-2001; 2001WO-US29871.  
XX  
XX 25-SEP-2000; 2000US-234925P.  
XX 12-JAN-2001; 2001WO-US00911.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Ruben SM, Komatsu Jis G, Duan DR, Rosen CA, Moore PA, Shi Y;  
XX Lafleur DM, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;  
XX Endress GA, Mucenski M, Ebner R;  
XX  
XX WPI: 2002-362489/39.  
XX N-PSDB; ABR69743.  
XX  
XX Novel 71 isolated secreted polypeptides and polynucleotides encoding  
XX the polypeptides, useful for treating Huntington's disease, sepsis,  
XX meningitis, thrombocytopenia, haemolytic anaemia, rheumatoid arthritis,  
XX asthma  
XX  
XX Claim 11, Page 1231-1232; 1478pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (or its  
XX fragment, homologue complement or allelic variant) encoding a human  
XX secreted protein (and its fragment, domain, epitope, variant, secreted  
XX form and species variant). Also included are a recombinant vector  
XX comprising the nucleic acid, a recombinant host cell comprising the  
XX vector, an antibody against the secreted protein, a recombinant host cell  
XX that expresses the secreted protein and a method of identifying a binding  
XX partner of the secreted protein. The nucleic acid and protein are used to  
XX prevent, diagnose, treat or ameliorate a medical condition in e.g.  
XX humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep  
XX for example autoimmune diseases e.g. rheumatoid arthritis,  
XX hyperproliferative disorders e.g. neoplasms of the breast or liver,  
XX cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
XX e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.  
XX Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
XX ocular disorders e.g. corneal infection. Many other diseases and  
XX disorders are listed in the specification. The polypeptides can also be  
XX used to aid wound healing an epithelial cell proliferation, to prevent  
XX skin aging due to sunburn, to maintain organs before transplantation, for  
XX improving cell culture of primary tissues, to regenerate tissues and in  
XX chemotaxis. The polypeptides can also be used as a food additive or  
XX preservative to increase or decrease storage capabilities. The present  
XX sequence represents a novel human secreted protein of the invention.  
SQ Sequence 472 AA;  
Query Match 100.0%; Score 883; DB 23; Length 472;  
Best Local Similarity 100.0%; Pred. No. 2.2e-83;  
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKFLFAFGGVLHSLSCGKAICNGISKRTFEEIKKEIASCCGVAKAIINLAVYGAQ 60  
DB 1 MKFLFAFGGVLHSLSCGKAICNGISKRTFEEIKKEIASCCGVAKAIINLAVYGAQ 60  
QY 61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYQNLQDGLKLVHLEPVRIPIHMERGEESA 120  
DB 61 NRSYERLALVDYVGPRLSGSKNLEKAIQIMYQNLQDGLKLVHLEPVRIPIHMERGEESA 120

QY 121 VMLEPRHKAIIAGLSSIGNPPEGITAEVLYVTSFPELORPASEANGKIIVYNO 175  
DB 121 VMLEPRHKAIIAGLSSIGNPPEGITAEVLYVTSFPELORPASEANGKIIVYNO 175  
RESULT 9  
ID ABR47892 standard; Protein: 472 AA.  
XX ABR47892;  
XX AC ABR47892;  
XX XX  
XX 12-JUN-2003 (first entry)  
XX  
XX Human secreted protein, SEQ ID 783.  
XX  
XX Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;  
XX KM vulnerable; antiinflammatory; nootropic; neuroprotective;  
XX antiParkinsonian; gene therapy; human; cardiovascular disorder.  
XX  
XX Homo sapiens.  
XX WO200295010-A2.  
XX  
XX 28-NOV-2002.  
XX  
XX 19-MAR-2002; 2002WO-US09785.  
XX  
XX 21-MAR-2001; 2001US-277340P.  
XX 19-JUL-2001; 2001US-306171P.  
XX 13-NOV-2001; 2001US-331287P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
XX  
XX WPI: 2003-129429/12.  
XX  
XX Novel human secreted proteins, useful for detecting, preventing,  
XX diagnosing, prognosticating, treating and/or ameliorating  
XX cardiovascular disorders such as arrhythmia -  
XX  
XX Claim 13; SEQ ID 783; 1881pp; English.  
XX  
XX The present invention relates to novel human secreted proteins  
XX (ABR47892-ABR48145) and their coding sequences (ACC5034-ACC5056). The  
XX proteins and their coding sequences are useful for the preparation of a  
XX diagnostic or pharmaceutical composition for diagnosing or treating a  
XX cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,  
XX coronary arteriosclerosis and myocardial ischaemia), neural disorders,  
XX immune system disorders, muscular disorders, reproductive disorders,  
XX gastrointestinal disorders, pulmonary disorders, renal disorders,  
XX proliferative disorders and/or cancerous diseases and conditions, for  
XX wound healing and epithelial cell proliferation, to treat inflammation or  
XX infection, for treating thrombosis and arteriosclerosis, for treating or  
XX preventing neural damage which occurs in neuronal disorders or  
XX neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
XX disease, to enhance bone and periodontal regeneration and aid in tissue  
XX transplants or bone grafts, to prevent skin aging or hair loss, to  
XX stimulate growth and differentiation of hematopoietic cells and bone  
XX marrow cells when used in combination with other cytokines, to maintain  
XX organs before transplantation or for supporting cell culture of primary  
XX tissues, to increase or decrease differentiation or proliferation of  
XX embryonic stem cells, or to modulate mammalian characteristics or  
XX metabolism.  
XX Note: The sequence data for this patent was published in electronic  
XX format and is available from WIPD at  
XX ftp.wipo.int/pub/published\_pct\_sequences.  
SQ Sequence 472 AA;  
Query Match 100.0%; Score 883; DB 24; Length 472;  
Best Local Similarity 100.0%; Pred. No. 2.2e-83;

Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKELIFAFPGVHLLSLCSGKAICKNGISKRTEFEIEKEIASCGDVAKAIINLAVYGAQ 60  
 DB 1 MKELIFAFPGVHLLSLCSGKAICKNGISKRTEFEIEKEIASCGDVAKAIINLAVYGAQ 60  
 QY 61 NRSYERLALLVDVTPGPRLLSGSKNLEKAIQIMYONLQODGLEKVLBPVRIIPMERGESA 120  
 DB 61 NRSYERLALLVDVTPGPRLLSGSKNLEKAIQIMYONLQODGLEKVLBPVRIIPMERGESA 120  
 QY 121 VMLEPRHKTALITGLSSIGTPPEGITAEVLVYTSFPELQRRASEARGKIIVYNO 175  
 DB 121 VMLEPRHKTALITGLSSIGTPPEGITAEVLVYTSFPELQRRASEARGKIIVYNO 175

RESULT 10  
 ABR48106  
 ID ABR48106 standard; Protein; 472 AA.

AC ABR48106;  
 XX  
 XX 12-JUN-2003 (first entry)  
 DT  
 DE Human secreted protein, SEQ ID 997.  
 XX  
 KM Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cyostatic;  
 KM vulnerable; antiinflammatory; nocrotropic; neuroprotective;  
 KM antiparkinsonian; gene therapy; human; cardiovascular disorder.  
 OS  
 OS Homo sapiens.  
 PN WO200295010-A2.  
 PD 28-NOV-2002.  
 XX  
 XX 19-MAR-2002; 2002WO-US09785.  
 PF  
 XX 21-MAR-2001; 2001US-277340P.  
 PR 19-JUL-2001; 2001US-306171P.  
 PR 13-NOV-2001; 2001US-331287P.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Ruben SM;  
 PI WPI; 2003-129429/12.  
 DR  
 PT Novel human secreted proteins, useful for detecting, preventing,  
 PT diagnosing, prognosticating, treating and/or ameliorating  
 PT cardiovascular disorders such as arrhythmia -  
 XX  
 XX Claim 13; SEQ ID 997; 1881bp; English.

The present invention relates to novel human secreted proteins (ABR47633-ABR48145) and their coding sequences (ACC50344-ACC50856). The proteins and their coding sequences are useful for the preparation of a diagnostic or pharmaceutical composition for diagnosing or treating a cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary arteriosclerosis and myocardial ischemia), neural disorders, immune system disorders, muscular disorders, reproductive disorders, gastrointestinal disorders, pulmonary disorders, renal disorders, CC proliferative disorders and/or cancerous diseases and conditions, for CC wound healing and epithelial cell proliferation, to treat inflammation or CC infection, for treating thrombosis and arteriosclerosis, for treating or CC preventing neural damage which occurs in neuronal disorders or CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's CC disease, to enhance bone and periodontal regeneration and aid in tissue CC transplants or bone grafts, to prevent skin aging or hair loss, to CC stimulate growth and differentiation of haematopoietic cells and bone CC marrow cells when used in combination with other cytokines, to maintain CC organs before transplantation or for supporting cell culture of primary CC tissues, to increase or decrease differentiation or proliferation of CC embryonic stem cells, or to modulate mammalian characteristics or

CC metabolism.  
 CC Note: The sequence data for this patent was published in electronic  
 CC format and is available from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX

Sequence 472 AA;  
 Query Match 100.0%; Score 883; DB 24; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 2, 2e-83;  
 Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKELIFAFPGVHLLSLCSGKAICKNGISKRTEFEIEKEIASCGDVAKAIINLAVYGAQ 60  
 DB 1 MKELIFAFPGVHLLSLCSGKAICKNGISKRTEFEIEKEIASCGDVAKAIINLAVYGAQ 60  
 QY 61 NRSYERLALLVDVTPGPRLLSGSKNLEKAIQIMYONLQODGLEKVLBPVRIIPMERGESA 120  
 DB 61 NRSYERLALLVDVTPGPRLLSGSKNLEKAIQIMYONLQODGLEKVLBPVRIIPMERGESA 120  
 QY 121 VMLEPRHKTALITGLSSIGTPPEGITAEVLVYTSFPELQRRASEARGKIIVYNO 175  
 DB 121 VMLEPRHKTALITGLSSIGTPPEGITAEVLVYTSFPELQRRASEARGKIIVYNO 175

RESULT 11  
 ABR00152  
 ID ABR00152 standard; Protein; 472 AA.

AC ABR00152;  
 XX  
 XX 03-APR-2003 (first entry)  
 DT  
 DE Human gene 142 encoded secreted protein HRACJ35, SEQ ID NO.441.  
 XX  
 KM Human; secreted protein; digestive disorder; gastrointestinal disorder;  
 KM mouth; oesophagus; stomach; small intestine; large intestine; liver;  
 KM biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;  
 KM immune disorder; inflammation; infection; wound healing; drug screening;  
 KM chromosome identification; chromosome mapping; cytostatic; gene therapy;  
 KM antiinflammatory; immunosuppressive; vulnetary; chromosome 8q22.2.  
 OS  
 OS Homo sapiens.  
 PN WO200276488-A1.  
 PD 03-OCT-2002.  
 XX  
 XX 19-MAR-2002; 2002WO-US08276.  
 PF  
 XX 21-MAR-2001; 2001US-277340P.  
 PR 19-JUL-2001; 2001US-306171P.  
 PR 13-NOV-2001; 2001US-331287P.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Ruben SM;  
 PI WPI; 2003-029900/02.  
 DR N-PSDB; ABZ71331.  
 DR  
 PT New human secreted proteins and nucleic acids, useful for detecting,  
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating  
 PT e.g. gastrointestinal diseases and disorders, or cancers -  
 XX  
 XX Claim 13; Page 1029-1030; 1216bp; English.

ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted  
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.  
 CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The  
 CC invention also encompasses antibodies specific for the secreted proteins,  
 CC the use of the secreted proteins in drug screening, and recombinant  
 CC vectors and host cells comprising a nucleic acid of the invention. The  
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody

fragments specific for the secreted proteins, and modulators of protein activity are useful for diagnosing, treating, ameliorating or preventing digestive disorders. Such conditions include disorders of the mouth, oesophagus, stomach, small intestine, large intestine, liver, biliary tract and pancreas, and include cancers of these organs and tissues. The secreted and proteins, and their nucleic acids may also be used in the treatment of immune disorders, inflammation, infection, hyperproliferative disorders, and to promote wound healing. Nucleic acids of the invention may be used for chromosome identification, chromosome mapping, in gene therapy, for identifying individuals from minute biological samples, as hybridisation probes, and as molecular weight markers. The present sequence represents a human secreted protein of the invention.

Sequence 472 AA:

Query Match 100.0%; Score 883; DB 24; Length 472;  
Best Local Similarity 100.0%; Pred. No. 2.2e-83;  
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MKFLIFAFGGVHLTSLCSGKAICNGISKRTPEIKKEIASCGDVAKAIINLAVYGAQ 60  
1 MKFLIFAFGGVHLTSLCSGKAICNGISKRTPEIKKEIASCGDVAKAIINLAVYGAQ 60  
61 NRSYERLALLVDYTPRLSSGKNLEKAIQIMYQNLQODGLEKYLEVPRIPIHMERGES 120  
61 NRSYERLALLVDYTPRLSSGKNLEKAIQIMYQNLQODGLEKYLEVPRIPIHMERGES 120  
121 VMLERPIHKIAILGLGSSIGTPPEGITAEVLVVTSPFELQRRASEARGKIVVYQ 175  
121 VMLERPIHKIAILGLGSSIGTPPEGITAEVLVVTSPFELQRRASEARGKIVVYQ 175

#### RESULT 12

ABR00274 ABR00274 standard; Protein; 472 AA.

AC ABR00274;

DT 03-APR-2003 (first entry)

Human gene 142 encoded secreted protein HRACJ35, SEQ ID NO:563.

Human; secreted protein; digestive disorder; gastrointestinal disorder; mouth; oesophagus; stomach; small intestine; large intestine; liver; biliary tract; pancreas; cancer; tumour; hyperproliferative disorder; immune disorder; inflammation; infection; wound healing; drug screening; chromosome identification; chromosome mapping; cytostatic; gene therapy; antiinflammatory; immunosuppressive; vulnery; chromosome 8q22.2.

Homo sapiens.

WO200276488-A1.

03-OCT-2002.

19-MAR-2002; 2002WO-US08276.

21-MAR-2001; 2001US-277340P.

19-JUL-2001; 2001US-306171P.

13-NOV-2001; 2001US-331287P.

(HUMA-) HUMAN GENOME SCT INC.

Rosen CA, Ruben SM;

WPI; 2003-029900/02.

N-PSDB; ABZ71453.

New human secreted proteins and nucleic acids, useful for detecting, preventing, diagnosing, prognosticating, treating and/or ameliorating e.g. gastrointestinal diseases and disorders, or cancers -

Claim 13; Page 1108-1109; 1216pp; English.

ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted protein genes, and ABP00011-ABP00299 represent the proteins they encode. ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The invention also encompasses antibodies specific for the secreted proteins, the use of the secreted proteins in drug screening, and recombinant vectors and host cells comprising a nucleic acid of the invention. The secreted proteins, nucleic acids encoding them, antibodies or antibody fragments specific for the secreted proteins, and modulators of protein activity are useful for diagnosing, treating, ameliorating or preventing digestive disorders. Such conditions include disorders of the mouth, oesophagus, stomach, small intestine, large intestine, liver, biliary tract and pancreas, and include cancers of these organs and tissues. The secreted proteins and their nucleic acids may also be used in the treatment of immune disorders, inflammation, infection, hyperproliferative disorders, and to promote wound healing. Nucleic acids of the invention may be used for chromosome identification, chromosome mapping, in gene therapy, for identifying individuals from minute biological samples, as hybridisation probes, and as molecular weight markers. The present sequence represents a human secreted protein of the invention.

Sequence 472 AA:

Query Match 100.0%; Score 883; DB 24; Length 472;  
Best Local Similarity 100.0%; Pred. No. 2.2e-83;  
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MKFLIFAFGGVHLTSLCSGKAICNGISKRTPEIKKEIASCGDVAKAIINLAVYGAQ 60  
1 MKFLIFAFGGVHLTSLCSGKAICNGISKRTPEIKKEIASCGDVAKAIINLAVYGAQ 60  
61 NRSYERLALLVDYTPRLSSGKNLEKAIQIMYQNLQODGLEKYLEVPRIPIHMERGES 120  
61 NRSYERLALLVDYTPRLSSGKNLEKAIQIMYQNLQODGLEKYLEVPRIPIHMERGES 120  
121 VMLERPIHKIAILGLGSSIGTPPEGITAEVLVVTSPFELQRRASEARGKIVVYQ 175  
121 VMLERPIHKIAILGLGSSIGTPPEGITAEVLVVTSPFELQRRASEARGKIVVYQ 175

#### RESULT 13

AA87081 AA87081 standard; Protein; 473 AA.

AC AA87081;

DT 09-MAY-2000 (first entry)

Human secreted protein sequence SEQ ID NO:120.

Human; secreted protein; diagnosis; cytostatic; immunosuppressive; antiinflammatory; nootropic; neuroprotective; antiallergic; cancer; tumour; neurodegenerative disorder; developmental abnormality; allergy; foetal deficiency; blood disorder; immune system disorder; arthritis; autoimmune disease; hepatic disease; renal disease; inflammation; Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis; infection; AIDS; spinal cord injury; transplant rejection; diabetes; asthma; sepsis; acne; psoriasis; cardiovascular disorder; reproductive disorder; gastrointestinal disorder; respiratory disorder; metabolic disorder; food additive; preservative.

Homo sapiens.

WO200004140-A1.

27-JAN-2000.

14-JUL-1999; 99WO-US15849.

15-JUL-1998; 98US-0092921.

15-JUL-1998; 98US-0092922.

PR 15-JUL-1998; 98US-0092956.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y, Pei  
PI Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE;  
PI Mucenski M, Endress GA, Soppet DR;  
XX  
DR WPI; 2000-161128/14.  
DR N-PSDB; AA298034.  
XX  
PT New isolated human genes, useful for diagnosis and treatment of, e.g.  
PT cancers, neurological or blood disorders  
XX  
PS Claim 11; Page 383-385; 494pp; English.  
XX  
CC The polynucleotide sequences given in AA298017 to AA298108 encode the  
CC human secreted proteins given in AA87064 to AA87223. Human secreted  
CC protein can have activities based on the tissues and cells the genes are  
CC expressed in. Examples of activities include: cytostatic;  
CC immunosuppressive; antiinflammatory; nootropic; neuroprotective; and  
CC antiallergic. The polynucleotides and their corresponding secreted  
CC polypeptides are useful for preventing, treating or ameliorating medical  
CC conditions, e.g. by protein or gene therapy. Also pathological conditions  
CC can be diagnosed by determining the amount of the new polypeptides in a  
CC sample or by determining the presence of mutations in the new  
CC polynucleotides. Human secreted protein s and their polynucleotides can  
CC be used for developing products for the diagnosis or treatment of cancer,  
CC tumours, neurodegenerative disorders, developmental abnormalities and  
CC foetal deficiencies, blood disorders, diseases of the immune system,  
CC autoimmune diseases, hepatic and renal disease, inflammation,  
CC allergies, Alzheimer's disease, behavioural disorders, schizophrenia,  
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,  
CC transplant rejection, diabetes, asthma, acne, psoriasis,  
CC cardiovascular disorders, reproductive disorders, gastrointestinal  
CC disorders, respiratory disorders and metabolic disorders. The  
CC proteins or polynucleotides can also be used as food additives or  
CC preservatives. The proteins are also useful for identifying their  
CC binding partners. AA298008 to AA298016 and AA87063 are sequence used in  
CC the exemplification of the present invention.  
XX  
SQ Sequence 473 AA;  
XX  
Query Match 100.0%; Score 883; DB 21; Length 473;  
Best Local Similarity 100.0%; Pred. No. 2.2e-83;  
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKFLIFAFPGVHLLSLCSGKAICKNGISKRTPEEIKKEIASCGDVAKAIINLA VYGAQ 60  
DB 1 MKFLIFAFPGVHLLSLCSGKAICKNGISKRTPEEIKKEIASCGDVAKAIINLA VYGAQ 60  
QY 61 NRSYERLALVDVTGPPRLSGSKNLEKAIQIMYQNLQODGLEKHLERVRIPHMERGES 120  
DB 61 NRSYERLALVDVTGPPRLSGSKNLEKAIQIMYQNLQODGLEKHLERVRIPHMERGES 120  
QY 121 VMLEPRIHKTAIILGLSSIGTPPEGITAEVLVWTSFDELQRRASEARKIIVVNO 175  
DB 121 VMLEPRIHKTAIILGLSSIGTPPEGITAEVLVWTSFDELQRRASEARKIIVVNO 175  
RESULT 14  
ABG09405  
ID ABG09405 standard; Protein; 474 AA.  
AC- ABG09405;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #9396.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX

OS Homo sapiens.  
XX  
XX W0200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US08631.  
XX  
XX 31-MAR-2000; 2000US-0540217.  
XX 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Dermanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX N-PSDB; AA573592.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensic, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity  
XX  
XX Claim 20; SEQ ID NO 39764; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX (II). (II) is useful for generating antibodies against it, detecting or  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical  
XX imaging of sites expressing (II). (I) and (II) are useful for treating  
XX disorders involving aberrant protein expression or biological activity.  
XX The polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensic, gene mapping, identification of mutations in  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG00010-ABG30377 represent novel human  
XX diagnostic amino acid sequences of the invention.  
XX Note: The sequence data for this patent did not appear in the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pat\_sequences.  
XX  
SQ Sequence 474 AA;  
XX  
Query Match 100.0%; Score 883; DB 22; Length 474;  
Best Local Similarity 100.0%; Pred. No. 2.2e-83;  
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKFLIFAFPGVHLLSLCSGKAICKNGISKRTPEEIKKEIASCGDVAKAIINLA VYGAQ 60  
DB 1 MKFLIFAFPGVHLLSLCSGKAICKNGISKRTPEEIKKEIASCGDVAKAIINLA VYGAQ 60  
QY 61 NRSYERLALVDVTGPPRLSGSKNLEKAIQIMYQNLQODGLEKHLERVRIPHMERGES 120  
DB 61 NRSYERLALVDVTGPPRLSGSKNLEKAIQIMYQNLQODGLEKHLERVRIPHMERGES 120  
QY 121 VMLEPRIHKTAIILGLSSIGTPPEGITAEVLVWTSFDELQRRASEARKIIVVNO 175  
DB 121 VMLEPRIHKTAIILGLSSIGTPPEGITAEVLVWTSFDELQRRASEARKIIVVNO 175  
RESULT 15  
AAM25762  
ID AAM25762 standard; Protein; 481 AA.  
AC AAM25762;  
XX  
XX 16-OCT-2001 (first entry)  
XX

XX Human protein sequence SEQ ID NO:1277.  
DE Human, cancer; ulcer; HIV infection; human immunodeficiency virus;  
XX anti-inflamatory; antirheumatic; antiallergic; immunosuppressive;  
KW antibacterial; endocrine; cardiac; central nervous system; virucide;  
KW anti-HIV; fungicide; antimutagen; cardiovascular; antineoplastic; anaemia;  
KW antiaggregant; haemostatic; vulnery; antitumor; osteopathic; eczema;  
KW dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic;  
KW immunoprotective; antidepressant; nocotropic; antiparkinsonian; infection;  
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
KW antianginal; rheumatoid arthritis; septic shock; pancreatitis;  
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;  
KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
XX neurological disorder.  
XX Homo sapiens.  
XX WO200153455-A2.  
XX 26-JUL-2001.  
XX 22-DEC-2000; 2000WO-US35017.  
XX 23-DEC-1999; 99US-0471275.  
XX 21-JAN-2000; 2000US-0488725.  
XX 25-APR-2000; 2000US-0552317.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Dmanac RT;  
XX WPI; 2001-457603/49.  
XX N-PSDB; AAH99703.  
XX Isolated human polynucleotides encoding polypeptides, useful for the  
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -  
XX Claim 20; Page 265; 1217pp; English.  
XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
CC AAM25963. The proteins can have activities based on the tissues and  
CC cells they are expressed in, such as: antiinflammatory; antirheumatic;  
CC antiallergic; immunosuppressive; antibacterial; endocrine; cardiac;  
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;  
CC cardiovascular; antineoplastic; haemostatic; vulnery;  
CC antitumor; osteopathic; dermatological; antiallergic; antiaesthetic;  
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nocotropic;  
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
CC encoding them can be used in gene therapy, antisense therapy and vaccine  
CC production. The proteins and polynucleotides are useful for screening for  
CC agonists or antagonists of a protein and for the treatment and diagnosis  
CC of disorders associated with the activity of a protein e.g. inflammation,  
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,  
CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
CC neurological disorders.  
XX SQ Sequence 481 AA;

Query Match 100.0%; Score 883; DB 22; Length 481;  
Best Local Similarity 100.0%; Pred. No. 2.3e-83;  
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAFGGVHLLSLCSGKALICNGISKRTFEEIKELIASCGDAKAIINLAVYGAQ 60  
|||||

Db 9 MKFLIFAFGGVHLLSLCSGKALICNGISKRTFEEIKELIASCGDAKAIINLAVYGAQ 68  
QY 61 NRSYERLALLVDVTGPPRLSGSKNLEKAIQIMYQNLQDDGLEKHLPEVRIPMERGESA 120  
Db 69 NRSYERLALLVDVTGPPRLSGSKNLEKAIQIMYQNLQDDGLEKHLPEVRIPMERGESA 128  
QY 121 VMLPPIHKITAILGLSSSIGTPPEGITAEVLYVTSPELORRASEARGKIVVYQ 175  
Db 129 VMLPPIHKITAILGLSSSIGTPPEGITAEVLYVTSPELORRASEARGKIVVYQ 183

Search completed: December 22, 2003, 11:49:04  
Job time : 30.7527 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 22, 2003, 11:37:50 ; Search time 14.3354 Seconds  
(without alignments)  
1173.984 Million cell updates/sec

Title: US-09-745-763-36\_COPY\_1\_175

Perfect score: 883  
Sequence: 1 MKFLIFAFGFGVHLISLCSG.....FDEIQRRSEARKIVVYQ 175

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_76:.\*  
2: pir1:.\*  
3: pir2:.\*  
4: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	217	24.6	493	2 T46974	leucyl aminopeptid
2	163.5	18.6	467	2 G67564	aminopeptidase, pr
3	93.5	10.6	557	2 G72356	penicillin-binding
4	92.5	10.5	1295	2 T24587	hypothetical prote
5	92	10.4	920	2 F81725	cell division prot
6	89.5	10.1	215	2 G83658	general stress pro
7	87	9.9	339	2 S27738	hypothetical prote
8	85.5	9.7	388	2 A80103	probable acetylglu
9	84.5	9.6	326	2 T36497	probable branched-
10	84.5	9.6	765	2 S74598	hypothetical prote
11	82.5	9.3	350	2 A98976	hypothetical prote
12	82.5	9.3	350	2 G85822	hypothetical prote
13	82	9.3	821	1 S76330	endopeptidase Clp
14	82	9.3	864	2 T48474	hypothetical prote
15	81.5	9.2	162	2 D64577	hypothetical prote
16	81	9.2	809	2 G86603	UDP-N-acetylmuram
17	81	9.2	809	2 D72022	protein T21B10.3 [
18	81	9.2	1132	2 D88250	hypothetical prote
19	81	9.2	1170	2 T25043	hypothetical prote
20	80	9.1	1486	2 C64832	cell division prot
21	80	9.1	1486	2 G90754	kinesin-like cell
22	80	9.1	1486	2 B85618	hypothetical prote
23	79	8.9	428	1 D71155	probable methyl-ac
24	79	8.9	428	1 F69001	hypothetical prote
25	79	8.9	1488	2 A80615	cell division prot
26	78.5	8.9	284	2 F70024	diiminopimelate ep
27	78.5	8.9	369	1 F0ADM2	minor core protein
28	78.5	8.9	370	1 H27748	probable glycerol
29	78.5	8.9	1863	1 A58881	breast/ovarian can

30	78	8.8	672	2 G87386	TPR domain protein
31	78	8.8	1812	2 I49350	breast/ovarian can
32	77.5	8.8	368	1 F0ADM5	minor core protein
33	76.5	8.7	492	2 T26936	hypothetical prote
34	76.5	8.7	761	2 E82167	5-methyltetrahydro
35	76.5	8.7	845	1 S34027	replication licens
36	76	8.6	605	2 D81748	DNA gyrase, chain
37	76	8.6	1535	2 T49042	hypothetical prote
38	75.5	8.6	414	2 F64386	hypothetical prote
39	75.5	8.6	579	2 S37953	hypothetical prote
40	75.5	8.6	730	2 G75292	excinuclease ABC c
41	75	8.5	180	2 S39775	alpha-s2-casein fo
42	75	8.5	227	2 D45522	variant surface gl
43	75	8.5	250	2 T05100	ribosomal protein
44	75	8.5	450	2 B70318	chaperone Hsp100 - A
45	75	8.5	564	2 A34325	juvenile-hormone e

#### ALIGNMENTS

RESULT 1  
T46974  
leucyl aminopeptidase [EC 3.4.11.-] ES-62 precursor [validated] - nematode (Acanthocheil)  
C:Species: Acanthocheilonema viteae  
C>Date: 17-Mar-2000 #sequence\_revision 17-Mar-2000 #text\_change 01-Feb-2002  
C:Accession: T46974  
R:Harnett, W.; Houston, K.M.; Tate, E.; Garate, T.; Apfel, H.; Adam, R.; Haslam, S.M.; P.  
Mol. Biochem. Parasitol. 104, 11-23, 1999  
A>Title: Molecular cloning and demonstration of an aminopeptidase activity in a filarial  
A:Reference number: 224335; MUID:20055923; PMID:10589978  
A:Accession: T46974  
A:Status: preliminary; translated from GB/EMBL/DDB  
A:Molecule type: mRNA  
A:Residues: 1-493 <HAR>  
A:Cross-references: EMBL:AF077194; NID:G3386477; PIDN:AAC28365.1; PID:G3386478  
C:Function:  
A:Description: EC 3.4.11.- leucyl aminopeptidase [validated, MUID:20055923]  
C:Keywords: aminopeptidase

Query Match 24.6%; Score 217; DB 2; Length 493;  
Best Local Similarity 32.5%; Pred. No. 2.7e-11;  
Matches 55; Conservative 32; Mismatches 58; Indels 24; Gaps 5;  
QY 15 LSLCSGKAIC-----KNGISKRTFEIEKEERIASCGDVAKATINNAVGAQNRSTYERL 67  
DB 14 LTVVLGAALVDPDKTAPDKNYIQETFGKVAR-----LIQYITKGEVGLAYQWL 62  
QY 68 ALVDVTPGRLSGSKNLEKAIQIMWQNLQODGLEKVLSPV-RIPMERGEESAVMLEPR 126  
DB 63 SKLVDFGCHRWGSDSLSEKIAFLSESLKNDNPKVTEVPNLPVHVRGNDVEMLEPR 122  
QY 127 IHKTAIIGLSSISICTPPEGITAEVLVYTSFDEIQRRSEARKIVVYQ 175  
DB 123 NQRLNVALIG---GSEBASATGEVTVYLDLV--KDDVVRGKIVVTAQ 166  
RESULT 2  
G87564  
aminopeptidase, probable [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: G87564  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.H.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon.  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A>Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: G87564  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-467 <STO>

R.Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiraoka, Y. *Nucleic Acids Res.* 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
 A:Reference number: A83650; MUID:20512582; PMID:11058132  
 A:Accession: C83658  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-215 <STO>  
 A:Cross-references: GB:AP001507; GB:BA000004; NID:g10172612; PIDN:BA03786.1; GSPDB:GN00  
 A:Experimental source: strain C-125  
 C:Genetics:  
 A:Gene: ctc  
 C:Superfamily: 50S ribosomal protein L25

Query Match 10.1%; Score 89.5; DB 2; Length 215;  
 Best Local Similarity 28.0%; Pred. No. 1.5;  
 Matches 44; Conservative 23; Mismatches 57; Indels 33; Gaps 8;

QY 44 GDVAKAI-----INLAVG---KAQNSYERLALL--VDYTPRLSGSKNLEKAI--QIM 91  
 DB 15 GSIVTRKIRKQGYVAVVAVGNKTKSQPISVEAVDFLKTREVGRNGLISLEVEKTKQKQVM 74  
 QY 92 YQNTQODGL--EKVHLE-----PVRIPHWERGEESAVMLEPRIHKIAIILGL 135  
 DB 75 VHDIQMPKADYHIDFEVDMSSEIRANVPVRLTGARVSEGVLSQLMETITVRSLS 134  
 QY 136 GSSIGTPPEGITAEVLVVTSPFDELQRRASERAKIV 172  
 DB 135 PADI---PEITLDVSSLAIGDSIQIR--DVRGNVNP 166

RESULT 7  
 S27738  
 hypothetical protein D - *Thermus aquaticus*  
 C:Species: *Thermus aquaticus*  
 C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 18-Jun-1993  
 C:Accession: S27738  
 R:Ashby, M.K.; Bergquist, P.L.  
 submitted to the EMBL Data Library, July 1990  
 A:Description: Cloning and sequence of IS1000 from *Thermus thermophilus* HB8.  
 A:Reference number: S27735  
 A:Accession: S27738  
 A:Molecule type: DNA  
 A:Residues: 1-339 <ASH>  
 A:Cross-references: EMBL:M33159  
 A:Note: the source is designated as *Thermus thermophilus*  
 C:Genetics:  
 A:Start codon: CTG

Query Match 9.9%; Score 87; DB 2; Length 339;  
 Best Local Similarity 26.5%; Pred. No. 4.5;  
 Matches 40; Conservative 21; Mismatches 52; Indels 38; Gaps 6;

QY 10 GCVHLISCSGKAIKNGISKRTPEIKETI-----ASGDVAKAI--INLAVG 57  
 DB 6 GGLFLSLIMEG-----RTLEEVEVFKRGHGLVPKKELEDLAKALEAGLLTLE 55  
 QY 58 KAQRSEYERLALLVDYTPRLSGSKNLEKAIQIMYQNTQODGLEKVLHPVRIHWERGE 117  
 DB 56 KVEARLKEERKLRERMRLAG-----LSTPEGEREA--RAFLFAFRASIPGSGE 104  
 QY 118 ESAVMLEP-----RIHKIAIILGLSSIGTPP 143  
 DB 105 EARVLMPHLPEPSRVPEVYGAALALEKTPP 135

RESULT 8  
 AB0103  
 probable acetylglucosamine-6-phosphate deacetylase YPO0838 [imported] - *Yersinia pestis*  
 C:Species: *Yersinia pestis*  
 C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
 C:Accession: AB0103  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M.B.  
 demo-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,

Nature 413, 523-527, 2001  
 A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
 A:Reference number: AB0001; MUID:21470413; PMID:11586360  
 A:Accession: AB0103  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-388 <KIR>  
 A:Cross-references: GB:AL590842; PIDN:CAC89685.1; PID:g15978912; GSPDB:GN00175  
 C:Genetics:  
 A:Gene: YPO0838

Query Match 9.7%; Score 85.5; DB 2; Length 388;  
 Best Local Similarity 24.6%; Pred. No. 7.2;  
 Matches 43; Conservative 26; Mismatches 53; Indels 53; Gaps 7;

QY 11 GCVHLISCSGKAIKNGISKRTF--EIKKEISCGVAAAIINLAVYKAQRSEYERLA 68  
 DB 218 GVHLYNGMSGILHREPCCGAVLYHDMLELADGIVHVPVNNLAV---RMKGYRRIA 273  
 QY 69 LLYD-----TVGPRLSGSK-NLEKAIQIMYQNTQODGL 100  
 DB 274 LITDQKAGGLGEGRYLLGQNHITVROGEARTDDGSLAGSTCSIDQLRMNIGHAQ---- 329  
 QY 101 EKVHLEPVRIPHWERGEESAVMLEPRIHKIAIILGLSSIGTPPEGITAEVLVVT 155  
 DB 330 -----VPWE-----AVQMSAV-PAAYIGLSTLSIQGQAQASVMVMS 369

RESULT 9  
 T36497  
 probable branched-chain alpha keto acid dehydrogenase E1 beta chain - *Streptomyces coelic*  
 C:Species: *Streptomyces coelicolor*  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
 C:Accession: T36497  
 R:Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, July 1999  
 A:Reference number: Z21608  
 A:Accession: T36497  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-326 <SAU>  
 A:Cross-references: EMBL:AL096822; PIDN:CAB46939.1; GSPDB:GN00070; SCOEDB:SCGD3.17C  
 A:Experimental source: strain A3(2)  
 C:Genetics:  
 A:Gene: bkdB1, SCOEDB:SCGD3.17C  
 C:Superfamily: pyruvate dehydrogenase (lipoamide) beta chain

Query Match 9.6%; Score 84.5; DB 2; Length 326;  
 Best Local Similarity 20.9%; Pred. No. 7.1;  
 Matches 36; Conservative 30; Mismatches 73; Indels 33; Gaps 6;

QY 33 FEETKEISA-----SCGDVAKAIINLAVY---GKAQRSEYERLALLVDYTPRLSGSKN 83  
 DB 93 YDQIVTQLAKHAKRSLSKVMKPVVIRIPYGGGIGAVHSHSEPEALPAHVAAGLKVSPSN 152  
 QY 84 LEXAIQIMYQNTQODGLEKVLHPVRIHWERGEESAVMLEPRIHKIAIILGLSSI---- 139  
 DB 153 AADAYMMQQAIGDD--PVITYFEPKR-RYWDKAEVDKKAIPGPHTARVAREGDTLLAA 210  
 QY 140 -----GTPPEGITAEVLVVTSPFDELQRRASERAKIV 173  
 DB 211 YGPVVKCREVADAABEGRSLEVDLRSISPDFTIQASVETRLIYVH 262

RESULT 10  
 S74598  
 hypothetical protein s11040 - *Synechocystis* sp. (strain PCC 6803)  
 C:Species: *Synechocystis* sp.  
 A:Variety: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
 C:Accession: S74598  
 R:Kaneko, T.; Sato, S.; Kocani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
 o, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996  
 A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*  
 A:Accession: S74322; MUID:97061201; PMID:8905231  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-765 <KAN>  
 A:Cross-references: EMBL:D90900; GB:AB001339; NID:g1651768; PIDN:BA16750.1; PID:d101748  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 9.3%; Score 82.5; DB 2; Length 350;  
 Best Local Similarity 20.6%; Pred. No. 20;  
 Matches 45; Conservative 38; Mismatches 64; Indels 71; Gaps 11;

QY 4 LIFAFGCVHLLSLSCGKAIKNGISKR-----TTEIKERIASCGDVAKIINLAVY--- 56  
 Db 548 LVLSFGGVAFVLAFLGRVVDMLNGTLIMTDRYALGDVQIDVGLVENMNIYITQ 607  
 QY 57 -----GK---AQN---RSYERLALVDVTVGPRLSGSKNLEKA---IOIMYQNL 95  
 Db 608 LRPGGRISTTPNGKISVQVQNLTKOMSRBEFIV-----IDQSNVDKALMLIRVSEOM 662  
 QY 96 QODGL-EKVHLEPVRI-----PHERGEESAVMLEPRHKA 131  
 Db 663 REDPLWQEKI-LEPAALIGVDIDASKGIRLQVWIKTQAGQHPVGRFRRLRW-----KKA 716  
 QY 132 ILGSSIGTPEEGITAEVLVVTSPDELORASEARKK 169  
 Db 717 FELAGIALGAPQORIV-----YHHGQKSSHNGK 746

RESULT 11  
 A:Accession: A98976  
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene  
 A:Reference number: A98629; MUID:21156231; PMID:11258796  
 A:Accession: A98976  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-350 <HAY>  
 A:Cross-references: GB:BA000007; PIDN:BA36200.1; PID:g13362245; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RIMD 0509952  
 A:Genetics: EC62777

Query Match 9.3%; Score 82.5; DB 2; Length 350;  
 Best Local Similarity 23.0%; Pred. No. 12;  
 Matches 35; Conservative 24; Mismatches 32; Indels 61; Gaps 6;

QY 28 ISKTFEIKERIASCGDVAKIINLAVYGAQNSYERLALVDVTVGPRLSGSKNLEKA 87  
 Db 190 LSEVDLDEVRERMSLG---SYIALNLEGSSQDRFTS-----LSIAENL--- 230  
 QY 88 IOIMYQNLQODGLEKVALE---PVRIPIHMERGEESA-----VMLEPRHKA- 131  
 Db 231 -----IKISETDIPVIYVHGPKGEDAKRVLDVCNNVRLSLSPSKRSLA 278  
 QY 132 -----ILGSSIGTPEEGITAE 149  
 Db 275 IIKDAYIAITPDTSTILHMASAVNTPVAIYAD-310

RESULT 12  
 G85822  
 C:Species: *Escherichia coli* (strain O157:H7, substrain EDL8

C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 A:Accession: G85822  
 A:Status: N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,  
 Miller, L.; Grobbeck, E.W.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: G85822  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-350 <STO>  
 A:Cross-references: GB:AB005174; NID:g12516156; PIDN:AAG57043.1; GSPDB:GN00145; UMG:P.Z31;  
 A:Experimental source: strain O157:H7, substrain EDL933  
 A:Genetics: Z3137

Query Match 9.3%; Score 82.5; DB 2; Length 350;  
 Best Local Similarity 23.0%; Pred. No. 12;  
 Matches 35; Conservative 24; Mismatches 32; Indels 61; Gaps 6;

QY 28 ISKTFEIKERIASCGDVAKIINLAVYGAQNSYERLALVDVTVGPRLSGSKNLEKA 87  
 Db 190 LSEVDLDEVRERMSLG---SYIALNLEGSSQDRFTS-----LSIAENL--- 230  
 QY 88 IOIMYQNLQODGLEKVALE---PVRIPIHMERGEESA-----VMLEPRHKA- 131  
 Db 231 -----IKISETDIPVIYVHGPKGEDAKRVLDVCNNVRLSLSPSKRSLA 278  
 QY 132 -----ILGSSIGTPEEGITAE 149  
 Db 279 IIKDAYIAITPDTSTILHMASAVNTPVAIYAD 310

RESULT 13  
 S76330  
 A:Accession: S76330  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S76330  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-821 <KAN>  
 A:Cross-references: EMBL:D64000; GB:AB001339; NID:g1001484; PIDN:BA110182.1; PID:g100155  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
 A:Function: allows c1p to hydrolyze polypeptides and proteins, probably by a chaperon  
 e activity. ATP hydrolysis is required for Clp hydrolysis of proteins but not of smaller  
 C:Superfamily: endopeptidase Clp ATP-binding chain  
 C:Keywords: ATP; hydrolase; molecular chaperone; nucleotide binding; P-loop; serine prot  
 F:208-215/Region: nucleotide-binding motif A (P-loop)  
 F:275-280/Region: nucleotide-binding motif B  
 F:545-552/Region: nucleotide-binding motif A (P-loop)  
 F:613-618/Region: nucleotide-binding motif B  
 F:714/Binding site: ATP (Lys) #status predicted  
 F:751/Binding site: ATP (Lys) #status predicted

Query Match 9.3%; Score 82; DB 1; Length 821;  
 Best Local Similarity 21.3%; Pred. No. 37;  
 Matches 43; Conservative 27; Mismatches 70; Indels 62; Gaps 6;

QY 19 SKAICNGISKTFEIKERIASCGDVAKIINLAVYGAQNSYERLALVDVTVGPR 78  
 Db 459 SKKRGNGEPEVTSEIATIVASWTCV-----VNLTSESESKLHMETLHQR 510

Qy	79	SGSNLEKAIQIMQNLC-----QDGEKXHLERVRIPIHREBESAV	121
Db	511	IGQEDAVAVAPRAIRPARVGLKNPNRPPIASFIFSGPTVGKTELTKALAAYPFGSEAMI	570
Qy	122	MLE-----PRIKRIALLGSSIGTPP-----EGITAEVLVYTSPELOR---	161
Db	571	RLDMSSEMERHTVSKL-----IGSPGYGVINEGQOLTEAVRRPPTVVLFEIERAMPD	625
Qy	162	-----RASEARGKIV	171
Db	626	IFNNLLQILLEDGRKLTDAKGRIV	647

RESULT 14

hypothetical protein T1E3.140 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T1E8474  
R:Beran, M.; Terry, N.; Ardiles, W.; Buyschaert, C.; Dasseville, R.; De Clerck, R.; De  
ewes, H.W.; Rüd, S.; Lemcke, K.; Mayer, K.F.X.  
Submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24491  
A:Accession: T1E8474  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-864 <BBV>  
A:Cross-references: EMBL:AL162972  
A:Experimental source: cultivar Columbia; BAC clone T1E3  
C:Genetics:  
A:Map position: 5  
A:Introns: 612/2; 626/3; 677/1; 696/3; 770/3  
A:Note: T1E3.140

### RESULT 15

Query Match	27.3%	Score 82	DB 2	Length 864
Best Local Similarity	1.1%	Pred. No. 39		
Matches	26	Conservative	23	Mismatches 29; Indels 18; Gaps 7

  

QY	13	HLHLGSG----	AICKNGISKRT-	BEIEETLSCGDVAKAIINLVYGAQNR	SYER	66
Db	267	MLALIEKQNHAYICSG	FGSNFVASSAVDM	YAKCSLRE	ST---	IFSEVDEKLE-
QY	67	LALLVDVGRPLSG-	SKNLE-KAIOIMYQNL	QDGL		100
Db	323	---LMTIT---	ISGPAKHAREVIL	LEKQDDDM		352

## A; CPOB-1

hypothetical protein HP0460 - Helicobacter pylori (strain 26695)  
C/Species: Helicobacter pylori  
C/Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C/Accession: D64577  
R/Tomb, J.F.: White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodet, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Meldrum, J.M.; Fujii, C.; Bowman, C.; Wathey, L.; Natore, J.D.; Kelley, J.M.; Cotton, M.D.; Meldrum, J.M.; Fujii, C.; Bowman, C.; Wathey, L. Nature 388, 539-547, 1997  
A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpik, P.D.; Smith, H.O.; Fraser, C. A/Titles: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A/Reference number: A64520; MUID:97394467; PMID:9252185  
A/Accession: D64577  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-162 <TOM>  
A/Cross-references: GB:AE000561; GB:AE000511; NID:g2313564; PIDN:AD07537.1; PID:g2313564

### Query Match

Query Match	9.2%	Score	81.5	DB	2	Length	163
Best Local Similarity	23.5%	Pred.	No. 5.5				
Matches	35	Conservative	31	Mismatches	60	Indels	23
						Gaps	6

QY 28 ISKRTEEIKEIASCGVAKAI INLAVYGKAQNRSYERLALVDIVGPRLSGSKNLEKA 87

[illegible]

Search completed: December 22, 2003, 11:53:01  
Job time : 16.3354 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: December 22, 2003, 11:34:20 ; Search time 11.0896 Seconds  
(without alignments)  
742.104 Million cell updates/sec

Title: US-09-745-763-36\_COPY\_1\_175  
Perfect score: 883  
Sequence: 1 MKFLIFAFSGVHLISCSG.....FDELQRRASBARKIVVYQ 175

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84.5	9.6	552	VNS1_EHDV2	P27585 epizootic h
2	81	9.2	809	MUDD_CHLPN	Q94701 c murc/dcl
3	80	9.1	1486	MUKB_ECOLI	P25523 escherichia
4	78.5	8.9	284	DAPE_BACSU	O32114 bacillus su
5	78.5	8.9	369	VCOM_ADEB2	P03267 human adeno
6	78.5	8.9	370	EGSA_AERPE	O94922 aeropyrum p
7	78.5	8.9	1863	BRCl_HUMAN	P38398 homo sapien
8	78	8.8	1812	BRCl_MOUSE	P48754 mus musculu
9	77.5	8.7	738	VCOM_ADEB5	P24938 human adeno
10	76.5	8.7	761	AB10_HUMAN	O94918 homo sapien
11	76.5	8.7	845	CC47_YEAST	P38132 saccharomyc
12	76.5	8.7	845	CC47_YEAST	P38132 saccharomyc
13	75.5	8.6	414	Y694_METUA	Q58105 methanococ
14	75.5	8.6	579	Y694_YEAST	P33343 saccharomyc
15	75.5	8.6	1286	SMC4_MOUSE	O8C947 mus musculu
16	75	8.5	180	CAS2_RABIT	P50418 corycolagus
17	75	8.5	450	HSIU_AQUAE	O65574 aquifex aeo
18	75	8.5	804	STL_BACSU	P36430 bacillus su
19	75	8.5	807	MCM3_XENLA	P47339 xenopus lae
20	75	8.5	1510	MUKB_HAEIN	P45187 haemophilus
21	74.5	8.4	183	PYRR_DEIRA	O94919 deinoococcus
22	74.5	8.4	1967	VG50_YEAST	P53327 saccharomyc
23	74	8.4	355	KIC3_HUMAN	O94960 homo sapien
24	74	8.4	390	SCC1_HUMAN	P29508 homo sapien
25	73.5	8.3	234	RS2_BACST	P81289 bacillus st
26	73	8.3	255	SOL4_YEAST	P53315 saccharomyc
27	73	8.3	344	Y613_METUA	Q58030 methanococ
28	73	8.3	428	CLPX_STRCO	O9E316 streptomyc
29	73	8.3	476	MRPB_NEUCR	P11913 neurospora
30	73	8.3	670	CKX1_ARATH	O30654 arabidopsis
31	73	8.3	893	MVP_HUMAN	O14764 homo sapien
32	72.5	8.2	205	RIVA_MYCBE	O49424 mycoplasma
33	72.5	8.2	351	DPO4_ECOS7	O84791 escherichia

34	72.5	8.2	351	1	DPO4_ECOL6	P59477 escherichia
35	72.5	8.2	441	1	ARCD_SCHPO	O74548 schizosacch
36	72.5	8.2	510	1	PURI_CHICK	P28173 gallus gall
37	72.5	8.2	881	1	YE8T_HAEIN	P44288 haemophilus
38	72	8.2	249	1	RS6_ONCMY	O94952 oncorhynch
39	72	8.2	429	1	YU64_AQUAE	O67776 aquifex aeo
40	72	8.2	886	1	ODPI_ECOLI	P06958 escherichia
41	71.5	8.1	223	1	GIDE_CORGL	O84153 corynebacte
42	71.5	8.1	303	1	CHEV_BACSU	P37599 bacillus su
43	71.5	8.1	411	1	MANA_CAEEL	P34650 caenorhabdi
44	71.5	8.1	513	1	CP12_MOUSE	P00186 mus musculu
45	71.5	8.1	1288	1	SMC4_HUMAN	Q94913 homo sapien

ALIGNMENTS

RESULT 1	VNS1_EHDV2	STANDARD;	PRT;	552 AA.
ID	VNS1_EHDV2			
AC	P27585;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	01-DEC-1992 (Rel. 24, Last annotation update)			
DE	Nonstructural protein NS1 (Hydrophobic tubular protein).			
GN	S6.			
OS	Epizootic hemorrhagic disease virus (serotype 2 / strain Alberta)			
OS	(Ehdv-2).			
OC	Viruses; dsRNA viruses; Reoviridae; Orbivirus.			
OX	NCBI_TaxID=10910;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=91253267; PubMed=1645906;			
RM	Nel L.H., Picard L.A., Huismans H.;			
RT	"A characterization of the nonstructural protein from which the			
RT	virus-specified tubules in epizootic haemorrhagic disease			
RT	virus-infected cells are composed."			
RL	Virus Res. 18:219-230(1991).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; M69085; AAA43000.1; -			
DR	EMBL; X59000; CAA41746.1; -			
DR	InterPro: IPR002630; Orbi NS1.			
DR	Pfam: PF01718; Orbi NS1; I.			
KW	Nonstructural protein.			
SQ	SEQUENCE 552 AA; 64553 MW; 6D4407083276B76F CRC64;			
Query Match	9.6%; Score 84.5; DB 1; Length 552;			
Best local Similarity	25.0%; Pred. No. 4.1;			
Matches	26; Conservative 19; Mismatches 46; Indels 19; Gaps 4;			
Qy	8 FFGVHLISCSGKAIKNGISKRTPEEI--KEEIASCGDVAKAIINLAVYKAQNRSY 64			
Db	329 FMOGRHRCQCL---FLKNGCDRETFYHIDVRTSEIWCSTVQVMIG-----EH 374			
Qy	65 ERLALVDVTPGRPLSGSKNEKAIQIMYQNLQDDGKVLDEPRIPHWERG 116			
Db	375 VDISLPQKLI--KLITGTEHLGRASDHFKYNAATGMEALIRTAIQIHRWIRG 424			
RESULT 2	MUDD_CHLPN	STANDARD;	PRT;	809 AA.
ID	MUDD_CHLPN			
AC	Q92701; Q9J085;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			

DR 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Murc/ddl bfunctional enzyme [includes: UDP-N-acetylmutamoyl-L-  
 DE alanine ligase (EC 6.3.2.8) (UDP-N-acetylmutamoyl-L-alanine  
 DE synthetase); D-alanine--D-alanine ligase (EC 6.3.2.4) (D-alanylanine  
 DE synthetase) (D-Ala-D-Ala ligase)].  
 DR MURC/DDL OR CPN0905 OR CP0961.  
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydiales.  
 RX NCBI\_TaxID=83558;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CW1029;  
 RX MEDLINE=99206606; PubMed=10192388;  
 RA Kalman S., Mitchell W., Marathe R., Jammal C., Fan J., Hyman R.W.,  
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;  
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";  
 RN Nat. Genet. 21:385-389(1999).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AR39;  
 RX MEDLINE=20150255; PubMed=10684935;  
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
 RA White O., Hickey E.K., Peterson J., Ullrich T., Berry K., Bass S.,  
 RA Liner K., Weidman J., Kouri H., Craven B., Bowman C., Dodson R.,  
 RA Gwin M., Nelson W., Deboy R., Kolony J., McClarty G., Salzberg S.L.,  
 RA Eisen J., Fraser C.M.;  
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia  
 RT pneumoniae AR39.";  
 RN Nucleic Acids Res. 28:1397-1406(2000).  
 RL [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J138;  
 RX MEDLINE=20330349; PubMed=10871362;  
 RA Shitai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
 RA Shiba T., Ishi K., Hattori M., Kuhara S., Nakazawa T.;  
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138  
 RT from Japan and CW1029 from USA.";  
 RL Nucleic Acids Res. 28:2311-2314(2000).  
 CC - FUNCTION: CELL WALL FORMATION.  
 CC - CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl + L-alanine = ADP +  
 CC phosphate + UDP-N-acetylmuramoyl-L-alanine.  
 CC - CATALYTIC ACTIVITY: ATP + 2 D-alanine = ADP + phosphate + D-  
 CC alanyl-D-alanine.  
 CC - PATHWAY: Peptidoglycan biosynthesis.  
 CC - SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC - SIMILARITY: In the C-terminal section; belongs to the murCDF  
 CC family.  
 CC - SIMILARITY: In the C-terminal section; belongs to the murCDF  
 CC family.  
 CC - SIMILARITY: In the C-terminal section; belongs to the D-alanine--  
 CC D-alanine ligase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL, AB001670; AAD19043.1; -;  
 DR EMBL, AB002254; AAF38741.1; -;  
 DR EMBL, AP002548; BAA99113.1; -;  
 DR PIR, D72022; D72022.  
 DR PIR, G86603; G86603.  
 DR HSSP, P07862; 110W.  
 DR PHCI-ZDPAGE; Q92701; -;  
 DR TIGR; CP0961; -;  
 DR HAMAP; MF\_00046; fused; 1.  
 DR HAMAP; MF\_00047; fused; 1.  
 DR InterPro; IPR005905; D\_ala\_D\_ala.  
 DR InterPro; IPR000291; Data\_1ig\_Van.  
 DR InterPro; IPR000713; Mur\_1igase.  
 DR InterPro; IPR004101; Mur\_1igase\_C.  
 DR InterPro; IPR005758; MurC.

DR Pfam; PF01820; Data Data ligase; 1.  
 DR Pfam; PF01825; Mur\_1igase; 1.  
 DR Pfam; PF02875; Mur\_1igase\_C; 1.  
 DR TIGRFAMs; TIGR01205; D\_ala\_D\_alatigr; 1.  
 DR TIGRFAMs; TIGR01082; murC; 1.  
 DR PROSITE; PS00843; DATA DATA LIGASE 1; 1.  
 DR PROSITE; PS00844; DATA DATA LIGASE 2; FALSE NEG.  
 KW Peptidoglycan synthetase; Cell wall; Cell division; Ligase;  
 KW ATP-binding; Multifunctional enzyme; Complete proteome.  
 FT DOMAIN 1 450  
 FT DOMAIN 451 809  
 FT NP\_BIND 111 117  
 FT ATP BIND 117  
 SQ SEQUENCE 809 AA; 89963 MW; 1A204C6B20E03B47 CRC64;  
 Query March 9.2%; Score 81; DB 1; Length 809;  
 Best Local Similarity 24.2%; Pred. No. 13;  
 Matches 48; Conservative 28; Mismatches 62; Indels 60; Gaps 9;  
 QY 17 LCSGAICKNGISRTFEIEKEBIA-SCGVAKAIINL-----AVYGAQ 60  
 DB 457 VCGGKS-CEHDSLSAQHVSQKYSPEFDVSYTIINRQGLMRTGKDPHPIERTGDSF 515  
 QY 61 NRSYERLALL-VDTVGPRLSGSKNLEKAIQIMQNLQOD----- 98  
 DB 516 LSSEIASALAKVDCLPVLGHPFGEDGTIOGFFELKPYAGPSLSAATAMDRLTKRI 575  
 QY 99 -----GLEKHLPEPRIRHMERGESAV--MLE-----PRIKIALIGSSIGTPPEGITA 148  
 DB 576 ASAVGVFVYPYQNPILCFWKNPELCTIQNLITPSPFMYKTA--HIGSSIG----- 625  
 QY 149 EVLVTSFDELORRASEA 166  
 DB 626 -IFLVRRKEELQEKISEA 642  
 RESULT 3  
 MURK\_ECOLI STANDARD; PRT; 1486 AA.  
 ID MURK\_ECOLI STANDARD; PRT; 1486 AA.  
 AC P22523; P71227; P77164; Q47398;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cell division protein murK.  
 GN MURK OR B0924.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / W3110;  
 RX MEDLINE=91114703; PubMed=1989883.  
 RA Niki H., Jaffe A., Imanura R., Ogura T., Hiraga S.;  
 RT "The new gene murK codes for a 177 kd protein with coiled-coil  
 RT domains involved in chromosome partitioning of E. coli.";  
 RL EMBO J. 10:183-193(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A. AND MUTANTS MURK33 AND MURK106.  
 RX MEDLINE=95080615; PubMed=7988894;  
 RA Yamanaka K., Mitani T., Feng Y., Ogura T., Niki H., Hiraga S.;  
 RT "Two mutant alleles of murK, a gene essential for chromosome  
 RT partition in Escherichia coli.";  
 RL FEBS Microbiol. Lett. 123:27-31(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).



[4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=97061202; PubMed=8905232;  
 RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,  
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horiuchi T.,  
 RT "A 718-Kb DNA sequence of the *Escherichia coli* K-12 genome  
 RT corresponding to the 12.7-28.0 min region on the linkage map.";  
 RL DNA Res. 3:137-155(1996).  
 RN [5]  
 RP SEQUENCE OF 1-44 FROM N.A.  
 RC STRAIN=K12 / W3110;  
 RX MEDLINE=94232180; PubMed=7513784;  
 RA Feng J., Yamana K., Niki H., Ogura T., Hiraga S.;  
 RT "New killing system controlled by two genes located immediately  
 RT upstream of the mukB gene in *Escherichia coli*.";  
 RL Mol. Genet. 243:136-147(1994).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1-227.  
 RX MEDLINE=20015369; PubMed=10545328;  
 RA van den Ent F., Lockhart A., Kendrick-Jones J., Loewe J.;  
 RT "Crystal structure of the N-terminal domain of MukB: a protein  
 RT involved in chromosome partitioning.";  
 RL Structure 7:1181-1187(1999).  
 CC -1- FUNCTION: ESSENTIAL FOR CHROMOSOME PARTITIONING. IMPLICATED IN  
 CC ATP-DEPENDENT CHROMOSOME PARTITIONING DURING CELL DIVISION.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

FT STRAND 18 23  
 FT HELIX 25 33  
 FT HELIX 35 51  
 FT TURN 55 57  
 FT HELIX 78 80  
 FT STRAND 81 81  
 FT STRAND 85 93  
 FT TURN 95 96  
 FT STRAND 99 108  
 FT TURN 112 114  
 FT STRAND 116 124  
 FT TURN 128 129  
 FT HELIX 132 136  
 FT STRAND 137 137  
 FT STRAND 146 147  
 FT HELIX 150 158  
 FT TURN 159 159  
 FT TURN 161 162  
 FT STRAND 164 167  
 FT HELIX 171 180  
 FT TURN 181 182  
 FT STRAND 184 184  
 FT HELIX 191 205  
 FT STRAND 210 220  
 FT STRAND 221 221  
 SQ SEQUENCE 1486 AA; 170229 MW; 38C7874BB78D6D6 CRC64;  
 Query Match 9.1%; Score 80; DB 1; Length 1486;  
 Best Local Similarity 27.0%; Pred. No. 34;  
 Matches 34; Conservative 17; Mismatches 33; Indels 42; Gaps 7;  
 QY 52 NLAIVYGAQRNRYVERLALVD-----TYGPRLSGSKN-LEKAIQIMYONLQDDGEXV 103  
 DB 1304 NRIITFSALAKLQRLNPQIDMCQRTPTTGEEBLDYRNLMEVEV---NRSGD--- 1356  
 QY 104 HLEPVRIPHERGESESAVMELEPRIHKIAIIGLSSISCTPEGITAELVVTSPDELQRA 163  
 DB 1357 -----WLRASG-----LSTGEAIGT---GMSILVMVQVMSDEBRR- 1391  
 QY 164 SEARGK 169  
 DB 1392 --LRGK 1395  
 RESULT 4  
 DAPF\_BACSU STANDARD; PRT; 284 AA.  
 AC 032114;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Diaminopimelate epimerase (EC 5.1.1.7) (DAP epimerase).  
 GN DAPF.  
 OS *Bacillus subtilis*.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID:1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borrill R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Broiliet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Emlich S.D., Emerson P.T.,  
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,  
 RA Gilm S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Guisepi G., Guy B.-J., Haga K., Halech U., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kaashara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Maul C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Odega S., Park S.H.,  
 RA Parro V., Pohl T.M., Portelle D., Porwolik S., Prescott A.M.,  
 RA Pireson E., Fujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Setiuchi J., Sekowaka A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpeira P., Tognoni A.,  
 RA Toledo V., Uchiyama S., Vandembol M., Vannier F., Vassarelli A.,  
 RA Viari A., Wambut R., Wedler E., Wedler H., Wellenreger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zunshtein E., Yoshikawa H., Zanchin A.,  
 RT "Type complete genome sequence of the Gram-positive bacterium *Bacillus*  
 RT *subtilis*.";  
 RL Nature 390:249-256(1997).  
 CC -1- CATALYTIC ACTIVITY: L1-2,6-diaminoheptanedioate = meso-  
 CC diaminoheptanedioate.  
 CC -1- PATHWAY: Biosynthesis of lysine from aspartate semialdehyde; sixth  
 CC step.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the diaminopimelate epimerase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z99120; CAB15207.1; -  
 DR PIR: F70024; F70024.  
 DR HSSP: P44859; 1BW2.  
 DR Subtilist: BG14048; dapF.  
 DR HAMAP: MF\_00197; -; 1.  
 DR InterPro: IPR001653; DAP epimerase.  
 DR Pfam: PF01678; DAP epimerase; 2.  
 DR Tricfams: TIGR00652; DapF; 1.  
 DR PROSITE: PS01326; DAP-EPIMERASE; 1.  
 DR KMW: Isomerase; Lysine biosynthesis; Complete proteome.  
 FT ACT SITE 76 76 BY SIMILARITY.  
 FT ACT SITE 226 226 BY SIMILARITY.  
 SQ SEQUENCE 284 AA; 30871 MW; 531D90C174C3BEC6 CRC64;  
 Query Match 8.9%; Score 78.5; DB 1; Length 284;  
 Best Local Similarity 18.8%; Pred. No. 6.4;  
 Matches 36; Conservative 31; Mismatches 85; Indels 39; Gaps 5;  
 QY 5 IFAPFGVHLSTLSCGKAICKNGISKRTPFEIKERIASCGVAKIINLAVYGRKNSY 64  
 DB 41 VYTGISDGLILICPSD---QAPVKKRIFFNDGSGKCGNLCRCVAKYAY---EHLGV 93  
 QY 65 ERLALLVDTVGPRLSGSKNLEKAIQIMYQNLQODLEKVLHPVRIPIHMERGEESAVMLE 124  
 DB 94 EETSLIET---LSGLVKAIVQVNGKYNVTVVMGERLRLKSLPLMDGEHEETIWT 149  
 QY 125 PRIHKAILGLSSIGTP-----PEGITAEVLVVTSPD 157  
 DB 150 MAFGEVELTGVAVSMGNPHIVPIADIEQAPLTTLGPVLEKDPREPPEGINAVEFETVNEQ 209  
 QY 158 ELQRASAEARG 168  
 DB 210 ELHFRVWE-RG 219  
 RESULT 5  
 VCOM\_ADR02 STANDARD; PRT; 369 AA.  
 AC P03267;  
 DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Minor core protein (Protein V).  
 GN PV.  
 OS Human adenovirus type 2.  
 CC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
 CC NCBI\_TaxID=10515;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85054835; PubMed=6094534;  
 RA Alestrom P., Ausjaerly G., Lager M., Yeh-Kai L., Pettersson U.,  
 RT J. Biol. Chem. 259:13980-13985(1984).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J01917; AAA92213.1; -  
 DR PIR: A03837; FOADM2.  
 DR InterPro: IPR005608; Adeno PV.  
 DR Pfam: PF03910; Adeno PV; 1.  
 KW Core protein; Late protein.  
 SQ SEQUENCE 369 AA; 41721 MW; 577C9B645B5E7DE CRC64;  
 Query Match 8.9%; Score 78.5; DB 1; Length 369;  
 Best Local Similarity 25.4%; Pred. No. 8.7;  
 Matches 47; Conservative 23; Mismatches 58; Indels 57; Gaps 9;  
 QY 29 SKRPFERKEIASCGV---AKIINLAVYGRKNSYERIALVDTVGPRLS----- 79  
 DB 90 STRTYKRVYDEVYDEDLLEQANERLGEFAVGK---RHKMLALPLDEGNPTPLKPVTL 146  
 QY 80 -----GSKNLEKAIQIMYQNLQODLEKVLHPVRIPIHMERGEES 119  
 DB 147 QQVPTLAPSEKRGKLRBSGDLAPVQVMVPRKQRLDEVLKRTVP-----GLEP 198  
 QY 120 AVMLEPRIHKAILGLG-----SSIGTPPEGITAEVLVVTSPD 164  
 DB 199 EVVRVP-IKQVA-FGLGVQTVVQIPTTSSTISATATEGMEQTQSPVAAVADAQVAAA 256  
 QY 165 EARCK 169  
 DB 257 AASK 261  
 RESULT 6  
 BGSA\_AERPE STANDARD; PRT; 370 AA.  
 ID BGSA\_AERPE  
 AC Q9YER2;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Glyceral-1-phosphate dehydrogenase [NAD(P)] (EC 1.1.1.261) (Sn-  
 DE glycerol-1-phosphate dehydrogenase) (G-1-P dehydrogenase)  
 DE (Enantiomeric glycerophosphate synthase).  
 GN BGSA OR APE0519.  
 OS Aeropyrum pernix.  
 CC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;  
 CC Desulfurococcaceae; Aeropyrum.  
 CC NCBI\_TaxID=56636;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KL;  
 RX MEDLINE=99310339; PubMed=10382966;  
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatawara Y.,  
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anket A., Kosugi H.,  
 RA Hasegawa A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,

RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,  
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,  
 RT "Complete genome sequence of an aerobic hyper-thermophilic  
 RT crenarchaeon, Aeropyrum pernix K1.";  
 RT DNA Res. 6:83-101(1999).  
 CC -1- FUNCTION: Responsible for the formation of archaea-specific  
 CC glycerophosphate backbone of phospholipids, G-1-P, from  
 CC dihydroxyacetonephosphate (DHAP) (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Sn-glycerol-1-phosphate + NAD(P)(+) =  
 CC glycerone phosphate + NAD(P)H.  
 CC -1- PATHWAY: De novo phospholipid biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).  
 CC -1- SIMILARITY: BELONGS TO THE GLYCEROL-1-PHOSPHATE DEHYDROGENASE  
 CC FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AP000059; BAA79484.1; -.  
 DR PIR; H72748; H72748.  
 DR HAMAP; MF\_00497; -; 1.  
 DR InterPro; IPR002658; DHQ\_synthase.  
 DR Pfam; PF01761; DHQ\_synthase, 1.  
 KW Phospholipid biosynthesis; Oxidoreductase; NADP; Complete proteome.  
 SQ SEQUENCE 370 AA; 39352 MW; 0782087EB9FCBF01 CRC64;  
 Query March 8.9%; Score 78.5; DB 1; Length 370;  
 Best Local Similarity 30.4%; Pred. No. 8.8;  
 Matches 42; Conservative 18; Mismatches 47; Indels 31; Gaps 7;  
 QY 55 YGKAQNRSYERLALVDVTGPRLSGSKULEKAIQIMYQNLQDGLKLEPVRIIPME 114  
 DB 46 VSGVAQGSY---LVVVS--GPTVS-SKYFER---LRASLEARGL-TVGKIRIDATVE 94  
 QY 115 RGEESA-VMEPRHKAIALGLSSIGTTP-----EITAEVLVVT 154  
 DB 95 TAEVVAEALAESRIEVAAGLGGKSIDVAKYASKRAGSVFVSIPTVASHDITSPFSLK 154  
 QY 155 SFDELQRRASEARGKIV 172  
 DB 155 GFDKPISRPAKAPRAIIT 172  
 RESULT 7  
 BRCA1\_HUMAN STANDARD; PRT; 1863 AA.  
 AC P38398;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Breast cancer type 1 susceptibility protein.  
 GN BRCA1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND VARIANT ARG-1775.  
 RX MEDLINE=95025896; PubMed=754954;  
 RA Miki Y., Swensen J., Shattuck-Bidens D., Futreal P.A., Harshman K.,  
 RA Tavtigian S., Liu Q., Cochran C., Bennett L.M., Ding W., Bell R.,  
 RA Rosenthal J., Husey C., Tran T., McClure M., Frye C., Hattler T.,  
 RA Phelps R., Haugen-Strano A., Katcher H., Yakumo K., Gholami Z.,  
 RA Shaffer D., Stone S., Bayer S., Wray C., Bogden R., Dayananth P.,  
 RA Ward J., Tonin P., Narod S., Bristow P.K., Norris F.H., Helvering L.,  
 RA Morrison P., Roestek P., Lai M., Barrett J.C., Lewis C., Neuhausen S.,  
 RA Cannon-Albright L., Goddard D., Wiseman R., Kamb A., Skolnick M.H.;  
 RT "A strong candidate for the breast and ovarian cancer susceptibility

RT gene BRCA1.";  
 RL Science 266:66-71(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97092865; PubMed=8938427;  
 RA Smith T.W., Lee M.K., Szabo C.I., Jerome N., McEuen M., Taylor M.,  
 RA Hood L., King M.-C.;  
 RT "Complete genomic sequence and analysis of 117 kb of human DNA  
 RT containing the gene BRCA1.";  
 RL Genome Res. 6:1029-1049(1996).  
 RN [3]  
 RP FUNCTION AS A E2-DEPENDENT UBIQUITIN LIGASE.  
 RX MEDLINE=99432238; PubMed=10500182;  
 RA Lorick K.L., Jensen J.P., Fang S., Ong A.M., Hatakeyama S.,  
 RA Weissman A.M.;  
 RT "RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent  
 RT ubiquitination.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:11364-11369(1999).  
 RN [4]  
 RP IDENTIFICATION OF BRCA1 AS MEMBER OF BASC.  
 RX MEDLINE=20245492; PubMed=10783165;  
 RA Wang Y., Cortez D., Yazdi P., Neff N., Elledge S.J., Qin J.;  
 RT "BASC, a super complex of BRCA1-associated proteins involved in the  
 RT recognition and repair of aberrant DNA structures.";  
 RL Genes Dev. 14:927-939(2000).  
 RN [5]  
 RP INTERACTION WITH SMC11.  
 RX MEDLINE=21866464; PubMed=11877377;  
 RA Yazdi P.T., Wang Y., Zhao S., Patel N., Lee E.Y.-H.P., Qin J.;  
 RT "SMC1 is a downstream effector in the ATM/ATR branch of the human  
 RT S-phase checkpoint.";  
 RL Genes Dev. 16:571-582(2002).  
 RN [6]  
 RP PHOSPHORYLATION BY ATM, AND MUTAGENESIS OF SER-1387, SER-1423 AND  
 RP SER-1524.  
 RX MEDLINE=22170551; PubMed=12183412;  
 RA Xu B., O'Donnell A.H., Kim S.-T., Kasten M.B.;  
 RT "Phosphorylation of serine 1387 in Brca1 is specifically required for  
 RT the ATM-mediated S-phase checkpoint after ionizing irradiation.";  
 RL Cancer Res. 62:4586-4591(2002).  
 RN [7]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=96400954; PubMed=8807330;  
 RA Couch F.J., Weber B.L.;  
 RT "Mutations and polymorphisms in the familial early-onset breast  
 RT cancer (BRCA1) gene. Breast Cancer Information Core.";  
 RL Hum. Mutat. 8:8-18(1996).  
 RN [8]  
 RP VARIANTS LEU-1637; GLU-1708 AND ARG-1775.  
 RX MEDLINE=95025878; PubMed=7939630;  
 RA Futreal P.A., Liu Q., Shattuck-Bidens D., Cochran C., Harshman K.,  
 RA Tavtigian S., Bennett L.M., Haugen-Strano A., Swensen J., Miki Y.,  
 RA Eddington K., McClure M., Frye C., Weaver-Felhaus J., Ding W.,  
 RA Gholami Z., Soederqvist P., Terry L., Jhanwar S., Berchuk A.,  
 RA Igglehart J.D., Marks J., Ballinger D.G., Barrett J.C., Skolnick M.H.,  
 RA Kamb A., Wiseman R.;  
 RT "BRCA1 mutations in primary breast and ovarian carcinomas.";  
 RL Science 266:120-122(1994).  
 RN [9]  
 RP VARIANTS BC GLY-64, AND VARIANTS ALA-772; ASN-1040 AND GLY-1443.  
 RX MEDLINE=95201806; PubMed=7894491;  
 RA Castilla L.H., Couch F.J., Erdos M.R., Hoekins K.F., Galzone K.,  
 RA Garber J.E., Boyd J., Lubin M.B., Deshano M.L., Brody L.C.,  
 RA Collins F.S., Weber B.L.;  
 RT "Mutations in the BRCA1 gene in families with early-onset breast and  
 RT ovarian cancer.";  
 RL Nat. Genet. 8:387-391(1994).  
 RN [10]  
 RP VARIANT BC GLY-61, AND VARIANTS ARG-356; GLY-1038; ASN-1040; ARG-1183  
 RP AND GLY-1613.  
 RX MEDLINE=95201808; PubMed=7894493;  
 RA Friedman L.S., Ostermeyer E.A., Szabo C.I., Dowd P., Lynch E.D.,  
 RA Rowell S.E., King M.-C.;

RT "Confirmation of BRCA1 by analysis of germline mutations linked to  
RT breast and ovarian cancer in ten families.";  
RL Nat. Genet. 8:399-404(1994).  
RN [11]  
RP VARIANT BC GLY-61.  
RX MEDLINE=96108965; PubMed=8554067;  
RA Serova O., Montagna M., Torchard D., Nared S.A., Tonin P., Sylia B.,  
RA Lynch H.T., Peunteun J., Lenoir G.M.;  
RT "A high incidence of BRCA1 mutations in 20 breast-ovarian cancer  
RT families.";  
RL Am. J. Hum. Genet. 58:42-51(1996).  
RN [12]  
RP VARIANT BOC TRP-841.  
RX MEDLINE=97123469; PubMed=8968716;  
RA Barker D.F., Almeida E.F.A., Casey G., Fain P.R., Liao S.-Y.,  
RA Maenaka I., Noble B., Kurosaki T., Anton-Culver H.;  
RT "BRCA1 R841W: a strong candidate for a common mutation with moderate  
RT phenotype.";  
RL Genet. Epidemiol. 13:595-604(1996).  
RN [13]  
RP VARIANTS BC AND BOC.  
RX MEDLINE=96372821; PubMed=8776600;  
RA Durocher F., Shattuck-Eidens D., McClure M., Labrie F.,  
RA Skolnick M.H., Goldgar D.E., Simard J.;  
RT "Comparison of BRCA1 polymorphisms, rare sequence variants and/or  
RT missense mutations in unaffected and breast/ovarian cancer  
RT populations.";  
RL Hum. Mol. Genet. 5:835-842(1996).  
RN [14]  
RP VARIANTS BC MET-271 AND SER-1150.  
RX MEDLINE=96303704; PubMed=8723683;  
RA Katagiri T., Eml M., Ito I., Kobayashi K., Yoshimoto M., Iwase T.,  
RA Kasumi F., Miki Y., Skolnick M.H., Nakamura Y.;  
RT "Mutations in the BRCA1 gene in Japanese breast cancer patients.";  
RL Hum. Mutat. 7:334-339(1996).  
RN [15]  
RP VARIANT BC GLY-61, AND VARIANTS ARG-239; TRP-841 AND ILE-1512.  
RX MEDLINE=98430998; PubMed=9760198;  
RA Dong J., Chang-Claude J., Wu Y., Schumacher V., Debatin I., Tonin P.,  
RA Royer-Pokora B.;  
RT "A high proportion of mutations in the BRCA1 gene in German  
RT breast/ovarian cancer families with clustering of mutations in the 3'  
RT third of the gene.";  
RL Hum. Genet. 103:154-161(1998).  
RN [16]  
RP VARIANT BC GLY-64, AND VARIANTS.  
RX MEDLINE=98141685; PubMed=9482581;  
RA Andersen T.I., Eiken H.G., Couch F., Kaada G., Skrede M., Johnsen H.,  
RA Aloysius T.A., Tveit K.M., Tranebjaerg L., Doerum A., Moeller P.,  
RA Weber B.L., Boerresen-Dale A.-L.;  
RT "Constant denaturant gel electrophoresis (CDGE) in BRCA1 mutation  
RT screening.";  
RL Hum. Mutat. 11:166-174(1998).  
RN [17]  
RP VARIANTS BC SER-22; LEU-461; ASP-465; VAL-552; SER-892; ASP-960;  
RX ILE-1025 AND ALA-1047.  
MEDLINE=98272917; PubMed=9609997;  
RA Katagiri T., Kasumi F., Yoshimoto M., Nomizu T., Aaishi K., Abe R.,  
RA Tsuchiya A., Sugano M., Takai S., Yoneda M., Fukutomi T., Namba K.,  
RA Makita M., Okazaki H., Hirata K., Okazaki M., Furutsuna Y.,  
RA Morishita Y., Iino Y., Karino T., Ayabe H., Hara S., Kajiwara T.,  
RA Housa S., Shimizu T., Toda M., Yamazaki Y., Uchida T., Kunikida K.,  
RA Sonoo H., Kurebayashi J.-I., Shimotsuna K., Nakamura Y., Miki Y.;  
RT "High proportion of missense mutations of the BRCA1 and BRCA2 genes in  
RT Japanese breast cancer families.";  
RL J. Hum. Genet. 43:42-48(1998).  
RN [18]  
RP VARIANT OC ARG-1749.  
RX MEDLINE=20455732; PubMed=10486320;  
RA Gayther S.A., Russell P., Harrington P., Antoniou A.C., Easton D.F.,  
RA Ponder B.A.J.;  
RT "The contribution of germline BRCA1 and BRCA2 mutations to familial  
RT ovarian cancer: no evidence for other ovarian cancer-susceptibility

RT genes.";  
RL Am. J. Hum. Genet. 65:1021-1029(1999).  
RN [19]  
RP VARIANT BC SER-346, AND VARIANTS LEU-871; GLY-1038; ARG-1183 AND  
RP GLY-1613.  
RX MEDLINE=99254821; PubMed=10323242;  
RA Li S.S.-L., Tseng H.-M., Yang T.-P., Liu C.-H., Teng S.-J.,  
RA Huang H.-W., Chen L.-M., Kao H.-W., Chen J.-H., Tseng J.-N., Chen A.,  
RA Hou M.-F., Huang T.-J., Chang H.-T., Mok K.-T., Tsai J.-H.;  
RT "Molecular characterization of germline mutations in the BRCA1 and  
RT BRCA2 genes from breast cancer families in Taiwan.";  
RL Hum. Genet. 104:201-204(1999).  
RN [20]  
RP VARIANTS OC, AND VARIANTS.  
RX MEDLINE=99214030; PubMed=10196379;  
RA Janzic S.A., Zlogas A., Krumroy L.M., Kraemer M., Plummer S.J.,  
RA Cohen P., Gildea M., Barker D., Halle R., Casey G., Anton-Culver H.;  
RT "Germline BRCA1 alterations in a population-based series of ovarian  
RT cancer cases.";  
RL Hum. Mol. Genet. 8:889-897(1999).  
CC -I- FUNCTION: Plays a central role in DNA repair by facilitating  
CC cellular response to DNA repair. Required for appropriate cell  
CC cycle arrests after ionizing irradiation in both the S-phase and  
CC the G2 phase of the cell cycle. Involved in transcriptional  
CC regulation of p21 in response to DNA damage. May function as a  
CC transcriptional regulator. Mediates E2-dependent ubiquitination.  
CC -I- SUBUNIT: Part of the BRCA1-associated genome surveillance complex  
CC (BASC), which contains BRCA1, MSH2, MSH6, MLH1, ATM, BLM, PMS2 and  
CC the RAD50-MRE11-NBS1 protein complex. This association could be a  
CC dynamic process changing throughout the cell cycle and within  
CC subnuclear domains. CtIP interacts specifically with the BRCT  
CC  
Query Match 8.9%; Score 78.5; DB 1; Length 1863;  
Best local similarity 22.8%; Pred. No. 62;  
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;  
QY 5 IFAPFGVHLISLCSGKA---ICNGISKRT-----FEERKEIAS--CG--DVAKAI 50  
DB 42 IFCKFCMLKLNQKGGSGQCELCNDITKXSLGSTSFQSLVEBLNITICAFOLDTGLEV 101  
QY 51 INLAVYGAQNRYSYERL---ALVDVTGVPRISSKNI-----EKAIQIYQNL-- 95  
DB 102 ANSYNFAKKNSSPEHLKDEVSIIQSGVYRNRAKRLQSPENPSLOETSLVSLNIGT 161  
QY 96 QQDLEKYNLEPVNIP--HMERGESAVMLEPRHKKIILIGSS--IGTPEGTAAVLV 152  
DB 162 VRLTRTKRQIQPKTSYIEIGSDS---EDTVNKATYCSVGQDELQITPPGTRDEI-- 216  
QY 153 VTSFDELQRRASE 165  
DB 217 --SLDSAKKAAACE 227  
RESULT 8  
BRCA1\_MOUSE STANDARD: PRT: 1812 AA.  
ID BRCA1\_MOUSE  
AC P48754; 060957; Q60983;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Breast cancer type 1 susceptibility protein homolog.  
GN BRCA1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Embryo;  
RX MEDLINE=96177659; PubMed=8634697;  
RA Abel K.J., Xy J., Yin G.Y., Lyons R.H., Weisler M.H., Weber B.L.;  
RT "Mouse Brca1: localization sequence analysis and identification of  
RT evolutionarily conserved domains.";

RL Hum. Mol. Genet. 4:2265-2273(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RX MEDLINE=96177660; PubMed=8634698;  
 RA Sharan S.K., Wims M., Bradley A.;  
 RT "Murine Brcal: sequence and significance for human missense  
 mutations.";  
 RL Hum. Mol. Genet. 4:2275-2278(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SvJ;  
 RX MEDLINE=96121367; PubMed=8575748;  
 RA Bennett L.M., Haugen-Strano A., Cochran C., Brownlee H.A.,  
 RA Fiedorek F.T. Jr., Wiseman R.W.;  
 RT "Isolation of the mouse homologue of BRCA1 and genetic mapping to  
 mouse chromosome 11.";  
 RL Genomics 29:576-581(1995).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SvJ; TISSUE=Embryo;  
 RX MEDLINE=96067162; PubMed=7590247;  
 RA Lane T.F., Deng C., Elson A., Lyu M.S., Kozak C.A., Leder P.;  
 RT "Expression of Brcal is associated with terminal differentiation of  
 ectodermally and mesodermally derived tissues in mice.";  
 RL Genes Dev. 9:2712-2722(1995).  
 RN [5]  
 RP SEQUENCE OF 727-1111 FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Embryo;  
 RX MEDLINE=96021028; PubMed=7550308;  
 RA Marquis S.T., Rajan J.V., Wynshaw-Boris A., Xu J., Yin G.Y.,  
 RA Abel K.J., Weber B.L., Chodosh L.A.;  
 RT "The developmental pattern of Brcal expression implies a role in  
 RT differentiation of the breast and other tissues.";  
 RL Nat. Genet. 11:17-26(1995).  
 RN [6]  
 RP SEQUENCE OF 789-1250 FROM N.A.  
 RC STRAIN=129/SvJ;  
 RX MEDLINE=96163506; PubMed=8566965;  
 RA Schroock E., Badger P., Larson D., Erdos M., Wynshaw-Boris A.,  
 RA Ried T., Brody L.;  
 RT "The murine homologue of the human breast and ovarian cancer  
 RT susceptibility gene Brcal maps to mouse chromosome 11D.";  
 RL Hum. Genet. 97:256-259(1996).  
 CC -1- FUNCTION: Plays a central role in DNA repair by facilitating  
 CC cellular response to DNA repair. Required for appropriate cell  
 CC cycle arrests after ionizing irradiation in both the S-phase and  
 CC the G2 phase of the cell cycle. Involved in transcriptional  
 CC regulation of p21 in response to DNA damage. May function as a  
 CC transcriptional regulator. Mediates E2-dependent ubiquitination (By  
 CC similarity).  
 CC -1- SUBUNIT: Part of the BRCA1-associated genome surveillance complex  
 CC (BASC), which contains BRCA1, MSH2, MSH6, MLH1, ATM, BLM, FMS2 and  
 CC the RAD50-MRE11-NBS1 protein complex. This association could be a  
 CC dynamic process changing throughout the cell cycle and within  
 CC subnuclear domains. CtIP interacts specifically with the BRCT  
 CC domains. Interacts with RNA polymerase II holoenzyme. Interacts  
 CC with SWC1L1 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -1- TISSUE SPECIFICITY: In the embryo, expressed in otic vesicles at  
 CC day 9.5. At day 10.5, this expression decreases and high levels  
 CC are found in the neuroectoderm. At days 11-12.5, high levels in  
 CC differentiating keratinocytes and whisker pad primordia. At days  
 CC 14-17, expression also observed in kidney epithelial cells. In  
 CC the adult, highest levels found in spleen, thymus, lymph nodes,  
 CC epithelial organs, and alveolar and ductal epithelial cells of  
 CC the mammary gland. Very low levels in brain, kidney, and skin. No  
 CC expression in heart, liver or lung.  
 CC -1- DEVELOPMENTAL STAGE: In the mammary gland, expression increases  
 CC dramatically during pregnancy. Levels fall during lactation and  
 CC increase again during post-lactational regression of the  
 CC mammary gland.  
 CC -1- PTM: Phosphorylated by ATM upon ionizing radiation (By  
 CC similarity).  
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
 CC -1- SIMILARITY: Contains 2 BRCT domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 CC EMBL; U35641; AAB17113.1; -;  
 CC EMBL; U31625; AAB17114.1; -;  
 CC EMBL; U32445; AAA963193.1; -;  
 CC EMBL; U36475; AAC52323.1; -;  
 CC EMBL; U33835; AAA97442.1; -;  
 CC MGI; MGI:104537; Brcal.  
 CC GO; GO:0005737; C:cytoplasm; IDA.  
 CC GO; GO:0008274; C:gamma-tubulin ring complex; ISS.  
 CC GO; GO:0005634; C:nucleus; ISS.  
 CC GO; GO:0016329; F:apoptosis regulator activity; ISS.  
 CC GO; GO:0005515; F:protein binding activity; ISS.  
 CC GO; GO:0016563; F:transcriptional activator activity; ISS.  
 CC GO; GO:0015631; F:tubulin binding activity; ISS.  
 CC GO; GO:0006915; P:apoptosis; ISS.  
 CC GO; GO:0007098; P:cell cycle; ISS.  
 CC GO; GO:0006978; P:DNA damage response, induction of cyclin-de. . .; ISS.  
 CC GO; GO:0045739; P:positive regulation of DNA repair; ISS.  
 CC GO; GO:0042127; P:regulation of cell proliferation; ISS.  
 CC GO; GO:0006559; P:regulation of transcription from Pol III pr. . .; ISS.  
 CC InterPro: IPR001357; BRCT.  
 CC InterPro: IPR002378; Brc1 cancer1.  
 CC InterPro: IPR001841; Znf\_Ting.  
 CC Pfam; PF00533; BRCT; 2.  
 CC Pfam; PF00097; ZF\_C3HC4; 1.  
 CC PRINTS; PR00493; BRCTCANCER1.  
 CC SMART; SM00292; BRCT; 2.  
 CC SMART; SM00184; RING; 1.  
 CC PROSITE; PSS0172; BRCT; 2.  
 CC PROSITE; PSS00518; ZF\_RING\_1; 1.  
 CC PROSITE; PSS0089; ZF\_RING\_2; 1.  
 CC DNA repair; Zinc-finger; DNA-binding; Nuclear protein; Polymorphism;  
 CC Anti-oncogene; Repeat; Phosphorylation.  
 CC ZN\_FING 24  
 CC FT 1585 1679  
 CC FT DOMAIN 1698 1797  
 CC FT DOMAIN 1562 1567  
 CC FT MOD\_RES 1343 1343  
 CC FT MOD\_RES 1481 1481  
 CC FT VARIANT 93 93  
 CC FT VARIANT 305 305  
 CC FT VARIANT 319 319  
 CC FT VARIANT 377 377  
 CC FT VARIANT 550 550  
 CC FT VARIANT 652 652  
 CC FT VARIANT 765 765  
 CC FT VARIANT 917 917  
 CC FT VARIANT 933 933  
 CC FT VARIANT 1122 1122  
 CC FT VARIANT 1206 1206  
 CC FT VARIANT 1212 1212  
 CC FT VARIANT 1255 1255  
 CC FT VARIANT 1261 1261  
 CC FT VARIANT 1264 1264  
 CC FT VARIANT 1269 1269  
 CC FT VARIANT 1283 1283  
 CC FT VARIANT 1337 1337  
 CC FT VARIANT 1349 1349  
 CC FT VARIANT 1352 1353  
 CC FT VARIANT 1381 1381  
 CC FT VARIANT 1390 1390  
 CC FT VARIANT 1400 1400  
 CC PHOSPHORYLATION (BY ATM) (BY SIMILARITY).  
 CC PHOSPHORYLATION (BY ATM) (BY SIMILARITY).  
 CC F -> L (IN STRAIN 129/SVJ).  
 CC T -> S (IN STRAIN 129/SVJ).  
 CC P -> A (IN STRAIN 129/SVJ).  
 CC Q -> E (IN STRAIN 129/SVJ).  
 CC K -> Q (IN STRAIN 129/SVJ).  
 CC A -> P (IN STRAIN 129/SVJ).  
 CC S -> P (IN STRAIN 129/SVJ).  
 CC P -> L (IN STRAIN 129/SVJ).  
 CC C -> S (IN STRAIN 129/SVJ).  
 CC K -> I (IN STRAIN 129/SVJ).  
 CC R -> R (IN STRAIN 129/SVJ).  
 CC RM -> GI (IN STRAIN 129/SVJ).  
 CC S -> R (IN STRAIN 129/SVJ).  
 CC H -> N (IN STRAIN 129/SVJ).  
 CC V -> A (IN STRAIN 129/SVJ).  
 CC P -> A (IN STRAIN 129/SVJ).  
 CC T -> K (IN STRAIN 129/SVJ).  
 CC T -> N (IN STRAIN 129/SVJ).  
 CC P -> T (IN STRAIN 129/SVJ).  
 CC EG -> OR (IN STRAIN 129/SVJ).  
 CC S -> P (IN STRAIN 129/SVJ).  
 CC G -> A (IN STRAIN 129/SVJ).  
 CC V -> D (IN STRAIN 129/SVJ).

FT VARIANT 1503 1503 E -> Q (IN STRAIN 129/SVJ).  
FT VARIANT 1549 1549 V -> A (IN STRAIN 129/SVJ).  
FT VARIANT 1680 1680 T -> K (IN STRAIN 129/SVJ).  
FT VARIANT 1712 1712 D -> E (IN STRAIN 129/SVJ).  
FT VARIANT 1721 1721 D -> E (IN STRAIN 129/SVJ).  
SQ SEQUENCE 1812 AA; 198669 MW; 2291EA74150BB86A CRC64;

Query Match 8.8%; Score 78; DB 1; Length 1812;  
Best Local Similarity 27.4%; Pred. No. 65;  
Matches 32; Conservative 22; Mismatches 37; Indels 26; Gaps 6;

QY 5 IFAFGVHLVSLCSGKA---ICXNGISKRT-----FEIKKEI-----ASCGDVAKAI 50  
DB 42 IFCKFCMLKLNQKKGSCPLCKMEITRSLQSGTRFSQALAEELRLMAAFELDTGMQL 101  
QY 51 INLVYGAQRNRYERL---ALLVTVG-----PRLS-GSKULEKAIQIMYONL 95  
DB 102 TNGFSFSKRRNNSCERLNEASIIQSGYRNRRLPQVEPGNATLKDSLGVOLSNL 158

## RESULT 9

VCOM\_ADE05 STANDARD; PRT; 368 AA.  
AC P24938;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE Minor core protein (Protein V).  
GN PV.  
OS Human adenovirus type 5.  
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
OX NCBI\_TaxID=28285;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92087470; PubMed=1727603;  
RA Chroboczek J., Bieber F., Jacrot B.;  
RT "The sequence of the genome of adenovirus type 5 and its comparison with the genome of adenovirus type 2.";  
RL Virology 186:280-285(1992).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M73260; AAA96409.1; -;  
DR PIR; C39449; FOADM5.  
DR InterPro; IPR005608; Adeno\_PV.  
DR Pfam; PF03910; Adeno\_PV.1.  
KW Core protein; Late protein.  
SQ SEQUENCE 368 AA; 41446 MW; 722E6C6D22C692A4 CRC64;

Query Match 8.8%; Score 77.5; DB 1; Length 368;  
Best Local Similarity 25.4%; Pred. No. 11;  
Matches 43; Conservative 21; Mismatches 50; Indels 55; Gaps 8;

QY 29 SKRTPEEKIEKIASCGDV---AKAIIINLVYGAQRNRYERLALLVTVGPRLS----- 79  
DB 89 STRTYKRYVDEYGDIEDLEQANERLGFAYCK---RKMDMALPLDSCGNTPPLSKPYTL 145  
QY 80 -----GSKULEKAIQIMYONL-OGLSEKVLHLEPVRIPIHMERGES 119  
DB 146 QOVLPALAPSEKRGKLRKESGDLAPTVQIMVPKRLDVLKMTVEP-----GHEP 197  
QY 120 AVMLEPRHIXIALILGSG-----SIGPPEGRITAEVLVYVS 155  
DB 198 EVRYVAP-IKOVA-PGLGVQTVDVQVLPITTSSTSIATGATGEMTQISPVAS 244

RESULT 10

AB10\_HUMAN STANDARD; PRT; 738 AA.  
ID AB10\_HUMAN  
AC Q9NR6G; Q13040; Q9H3V0.  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE ATP-binding cassette, sub-family B, member 10, mitochondrial precursor (ATP-binding cassette transporter 10) (ABC transporter 10 protein)  
DE (Mitochondrial ATP-binding cassette 2) (M-ABC2).  
DE ABCB10.  
GN Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymphoblast;  
RX MEDLINE=20382730; PubMed=10922475;  
RA Zhang F., Hogue D.L., Liu L., Fisher C.L., Hui D., Childs S., Ling V.;  
RT "M-ABC2, a new human mitochondrial ATP-binding cassette membrane protein.";  
RL FEBS Lett. 478:89-94(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Ito K., Suzuki H., Sugiyama Y.;  
RT Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 510-738 FROM N.A.  
RX MEDLINE=95284468; PubMed=7766993;  
RA Allkmeets R., Gerard B., Glavac D., Ravnik-Glavac M., Jenkins N.A., Gilbert D.J., Copeland N.G., Modi W., Dean M.;  
RT "Characterization and mapping of three new mammalian ATP-binding transporter genes from an EST database.";  
RL Mamm. Genome 6:114-117(1995).  
RN [4]  
RP VARIANT SSR-150.  
RX MEDLINE=21686803; PubMed=11829140;  
RA Saito S., Iida A., Sekine A., Mura Y., Ogawa C., Kawachi S., Higuchi S., Nakamura Y.;  
RT "Three hundred twenty-six genetic variations in genes encoding nine members of ATP-binding cassette, subfamily B (ABCB/MDR/TAP), in the Japanese population.";  
RL J. Hum. Genet. 47:38-50(2002).  
CC -1- FUNCTION: May mediate critical mitochondrial transport functions related to heme biosynthesis (By similarity).  
CC -1- SUBUNIT: Homodimer or heterodimer (Potential).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane.  
CC -1- TISSUE SPECIFICITY: Ubiquitous. Highly expressed in bone marrow, expressed at intermediate to high levels in skeletal muscle, small intestine, thyroid, heart, brain, placenta, liver, pancreas, prostate, testis, ovary, leukocyte, stomach, spinal cord, lymph node, trachea and adrenal gland, and low levels are found in lung, kidney, spleen, thymus and colon.  
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

[illegible]





ID YK04\_YEAST STANDARD; PRT; 579 AA.  
AC P32343;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hypothetical 65.1 kDa protein in RKN3-SRP21 intergenic region.  
GN YK1124W OR YK1529.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN  
RP SEQUENCE OF 1-120 FROM N.A.  
RA Rad M.R., Xu G., Kirchbach L., Fritz C., Keuchel H., Hollenberg C.P.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
RN  
RP SEQUENCE OF 41-579 FROM N.A.  
RX MEDLINE=9283952; PubMed=1514329;  
RA Colledge L., Richard G.-F., Thierry A., Dujon B.;  
RT "Sequence of a segment of Yeast chromosome XI identifies a new  
RT mitochondrial carrier, a new member of the G protein family, and a  
RT protein with the PAKK motif of the H1 histones."  
RL Yeast 8:325-336(1992).  
CC  
CC -1- SIMILARITY: TO YEAST YMR171C.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
CC EMBL: Z28123; CAAB1964.1; -  
DR EMBL: S44213; AAB23074.1; -  
DR PIR: S37953; S37953.  
DR SGD: S0001607; SSH4.  
DR InterPro: IPR003877; SPRY\_receptor.  
DR Pfam: PF00622; SPRY\_1.  
DR SMART: SM00449; SPRY\_1.  
KM Hypothetical protein.  
SQ SEQUENCE 579 AA; 65060 MM; 700FICBA07A9754 CRC64;  
SQ  
Query March 8.6%; Score 75.5; DB 1; Length 579;  
Best Local Similarity 24.1%; Pred. No. 28;  
Matches 32; Conservative 24; Mismatches 64; Indels 13; Gaps 4;  
QY 45 DVAKAI-----INLAVYGAQNRSYERIALIVD-----TWGPRLSGSKNLEKAIQIMYQNTL 95  
DB 321 DVTGNIGIDLTIGAGFAAATRTYTRGDLLEDDPNVSFRALBEGKDI EVAKDLQRVHD 380  
QY 96 QQDGLKXVLEPRVIRPWERGESAVMLKPRIRIKAIIGLSSIGTPPEGITAEV---LV 152  
DB 381 PHDESDEMTSDEVEL-HVNLGVQVFVFIEANVKYAFSGSVYGQIGIPAVNGTEIKKOTI 439  
QY 153 VTSFDELQRASE 165  
DB 440 LQKGELEPRYAD 452  
RESULT 15  
SMC4\_MOUSE  
ID SMC4\_MOUSE STANDARD; PRT; 1286 AA.  
AC Q8CG47; Q8BT57; Q8BT59; Q99K21;  
DT 15-SEP-2003 (Rel. 42, Created)  
DT 15-SEP-2003 (Rel. 42, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Structural maintenance of chromosomes 4-like 1 protein (Chromosome-  
DE associated polypeptide C) (XCAP-C homolog).  
GN SMC4L1 OR SMC4 OR CAPC.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;  
RN  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Cobbe N., Heck M.M.S.;  
RT "Phylogenetic analysis of SMC proteins."  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
RN  
RN [2]  
RP SEQUENCE OF 1-602 FROM N.A.  
RC SRRAIN=C57BI/657; and NOD: TISSUE=Embryonic liver, and Thymus;  
RX MEDLINE=22354683; PubMed=12466851;  
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
RA Nakaido I., Osato N., Saito R., Suzuki A., Yamataka I., Kiyosawa H.,  
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,  
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
RA Schirral L.M., Kanapin A., Matsuda H., Batalov S., Batzel K.W.,  
RA Blake J.A., Bradt D., Brasic V., Chochua C., Corbani L.E., Cousins S.,  
RA Dalla E., Dragani T.A., Fletcher C.P., Forrest A., Fraser K.S.,  
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavan W.J., Petrea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Saitana R., Schneider C., Semple C.A., Setton M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Velardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wyshaw-Boris A., Yamagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavalan M., Zhu Y., Zimmer C.J., Carninci P., Hayatsu N.,  
RA Hironaka-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yaunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
RN  
RN [3]  
RP SEQUENCE OF 1103-1286 FROM N.A.  
RC TISSUE=Breast tumor;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stedington M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.J.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywicki M.I., Skalska U., Smalilus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences".  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16993(2002).  
CC  
CC -1- FUNCTION: Central component of the condensin complex, a complex  
CC required for conversion of interphase chromatin into mitotic-like  
CC condensed chromosomes. The condensin complex probably introduces  
CC positive supercoils into relaxed DNA in the presence of type I  
CC topoisomerases and converts nicked DNA into positive knotted forms  
CC in the presence of type II topoisomerases (By similarity).  
CC  
CC -1- SUBUNIT: Forms a heterodimer with SMC2L1. Component of the  
CC condensin complex, which contains the SMC2L1 and SMC4L1  
CC heterodimer, and three non SMC subunits that probably regulate the  
CC complex: BRN1/CAP1, CNA1/CAP2 and CAP3 (By similarity).

CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic. In interphase  
CC cells, the majority of the condensin complex is found in the  
CC cytoplasm, while a minority of the complex is associated with  
CC chromatin. A subpopulation of the complex however remains  
CC associated with chromosome foci in interphase cells. During  
CC mitosis, most of the condensin complex is associated with the  
CC chromatin. At the onset of prophase, the regulatory subunits of  
CC the complex are phosphorylated by CDC2, leading to condensin's  
CC association with chromosome arms and to chromosome condensation.  
CC Dissociation from chromosomes is observed in late telophase (By  
CC similarity).  
CC -1- DOMAIN: The hinge domain, which separates the large intramolecular  
CC coiled coil regions, allows the heterodimerization with SMC2L1,  
CC forming a V-shaped heterodimer (By similarity).  
CC -1- SIMILARITY: Belongs to the SMC family. SMC4 subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL; AJ534940; CAD59183.1; -  
CC EMBL; AK088846; BAC40608.1; -  
CC EMBL; AK088350; BAC40297.1; -  
CC EMBL; BC005507; AA05507.1; -  
CC MGI; MGI:1917349; SMC4L1.  
CC InterPro; IPR003439; ABC\_transporter.  
CC InterPro; IPR003405; SMC\_C.  
CC InterPro; IPR003395; SMC\_N.  
CC Pfam; PF02483; SMC\_C; 1.  
CC Pfam; PF02463; SMC\_N; 1.  
CC DNA condensation; Mitosis; Cell cycle; ATP-binding; Coiled coil;  
CC Nuclear protein.  
CC NP\_BIND 111 118 ATP (POTENTIAL).  
CC FT DOMAIN 270 589 COILED COIL (POTENTIAL).  
CC FT DOMAIN 590 767 FLEXIBLE HINGE.  
CC FT DOMAIN 768 1018 COILED COIL (POTENTIAL).  
CC FT DOMAIN 1068 1133 COILED COIL (POTENTIAL).  
CC FT DOMAIN 1189 1224 ALA/ASP-RICH (DA-BOX).  
CC FT DOMAIN 64 69 POLY-PRO.  
CC SQ SEQUENCE 1286 AA; 146894 MW; CFEAD84199C3CEB5 CRC64;  
Query Match 8.6%; Score 75.5; DB 1; Length 1286;  
Best local Similarity 22.8%; Pred. No. 74; Mismatches 61; Indels 41; Gaps 12;  
Matches 41; Conservative 37; Mismatches 61; Indels 41; Gaps 12;  
QY 16 SLCSGKAIKNGISRTFEIEKEIASCGDAKAIINLAIVYGKAONRSYERIALLVDTVG 75  
Db 938 SVCRTEKIKD--TEKEINDKTEAKNIEDKAEVIN-----NTKTAETSL----- 981  
QY 76 PRLSGS-KULEKAIQIMYON--LQDGLK-KVHLEPV-----RIPHWERGEESAV 121  
Db 982 PEIQKEHNLQELKVIQENHALQKDALSIKLEQIDGHISENSKIKYWOK-EISKI 1040  
QY 122 MLEP----RIKRIALIGLSSIGTP-PEGITAEVLVTSFDELQRRASEAR-GKIVVYNQ 175  
Db 1041 KLHPVEDNPVETVAVLSQLQEELEAIRNPSSITNEIAL---EAQCREKMPNIGALAEYKK 1096

Search completed: December 22, 2003, 11:50:02  
Job time : 13.0896 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:37:00 ; Search time 26.507 Seconds  
(without alignments)  
1703.674 Million cell updates/sec

Title: US-09-745-763-36\_COPY\_1\_175  
Perfect score: 883  
Sequence: 1 MKFLIFAFGCVHLSLCSGKAIKNGISRTPEIEKEELIASGCDVAKAIINLAVYKQ 60

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	883	100.0	472.4	Q9Y646	Q9Y646 homo sapien
2	883	100.0	541.4	Q9Y5X6	Q9Y5X6 homo sapien
3	872	98.8	472.4	Q8NB21	Q8NB21 homo sapien
4	736	83.4	433.1	Q70216	Q70216 mus musculu
5	736	83.4	470.1	Q9WVJ3	Q9WVJ3 mus musculu
6	734	83.1	472.1	Q9JUV0	Q9JUV0 ratius norv
7	727	82.3	472.1	Q9Z1Y1	Q9Z1Y1 ratius norv
8	493	55.8	140.4	Q9UNM8	Q9UNM8 homo sapien
9	217	24.6	493.5	Q76552	Q76552 acanthochei
10	180.5	20.4	414.1	Q8EBH7	Q8EBH7 shewanella
11	163.5	18.5	467.1	Q9AS48	Q9AS48 caulobacter
12	121	13.7	472.1	Q8PH8	Q8PH8 xanthomonas
13	117	13.3	375.1	Q8P625	Q8P625 xanthomonas
14	93.5	10.6	557.1	Q9WZ59	Q9WZ59 thermotoga
15	92.5	10.5	1295.5	Q22257	Q22257 caenorhabdi
16	92	10.4	920.16	Q9PL78	Q9PL78 chlamydia m

17	91.5	10.4	731.2	Q8RT29	Q8RT29 pseudoalter
18	89.5	10.1	215.16	Q9KJ34	Q9KJ34 bacillus ha
19	89	10.1	592.6	Q9XSR3	Q9XSR3 canis famli
20	87	9.9	449.2	Q93EJ5	Q93EJ5 bacillus l1
21	86.5	9.8	770.17	Q8YPI	Q8YPI methanopyru
22	85.5	9.7	388.16	Q8ZHQ2	Q8ZHQ2 yersinia pe
23	85.5	9.7	729.2	Q9KH34	Q9KH34 antarctic b
24	85	9.6	349.2	Q8VQ34	Q8VQ34 spiroplasma
25	84.5	9.6	326.16	Q9XA61	Q9XA61 streptomyce
26	84.5	9.6	551.12	Q66601	Q66601 epizootic h
27	84.5	9.6	765.16	P72735	P72735 synchocyst
28	82.5	9.3	350.16	Q8X8V6	Q8X8V6 escherichia
29	82	9.3	524.4	Q9NKG3	Q9NKG3 homo sapien
30	82	9.3	593.16	Q8N152	Q8N152 homo sapien
31	82	9.3	821.16	Q55662	Q55662 synchocyst
32	82	9.3	864.10	Q9LZ19	Q9LZ19 arabidopsis
33	81.5	9.2	162.16	Q25207	Q25207 heliobacte
34	81.5	9.2	423.6	Q28014	Q28014 bradyphus tr
35	81.5	9.2	456.2	Q8KJH9	Q8KJH9 rhizobium e
36	81.5	9.2	551.12	Q66600	Q66600 epizootic h
37	81	9.2	579.10	Q941N0	Q941N0 hordium vul
38	81	9.2	1170.5	Q22624	Q22624 caenorhabdi
39	80	9.1	1479.2	Q937H2	Q937H2 klebsiella
40	80	9.1	1486.16	Q8XDG0	Q8XDG0 escherichia
41	80	9.1	1486.16	Q8FJA2	Q8FJA2 escherichia
42	79.5	9.0	377.16	Q8COD7	Q8COD7 staphylococ
43	79.5	9.0	409.6	Q95KU5	Q95KU5 cholepus d
44	79	8.9	428.17	Q58181	Q58181 pyrococcus
45	79	8.9	499.17	Q26204	Q26204 methanobact

## ALIGNMENTS

### RESULT 1

Q9Y646 PRELIMINARY; PRT; 472 AA.  
ID Q9Y646  
AC Q9Y646; 01-NOV-1999 (TREMURel. 12, Created)  
DT 01-NOV-1999 (TREMURel. 12, Last sequence update)  
DT 01-MAR-2003 (TREMURel. 23, Last annotation update)  
DE Aminopeptidase (Plasma glutamate carboxypeptidase).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Liu C.H., Lin B.Y., Chang L.Y.;  
RT "Cloning of the human aminopeptidase gene";  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RA Strausberg R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF107834; AAD43214.1; -;  
DR EMBL; BC020689; AAH20689.1; -;  
DR MEROPS; M28.014; -;  
DR InterPro; IPR00137; PA.  
DR PROSITE; PS0840; PA; 1.  
KW Carboxypeptidase.  
SQ SEQUENCE 472 AA; 51887 MW; EB6CBD2149E042BF CRC64;

Query Match 100.0%; Score 883; DB 4; Length 472;  
Best Local Similarity 100.0%; Pred. No. 1.9e-75;  
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLIFAFGCVHLSLCSGKAIKNGISRTPEIEKEELIASGCDVAKAIINLAVYKQ 60  
DB 1 MKFLIFAFGCVHLSLCSGKAIKNGISRTPEIEKEELIASGCDVAKAIINLAVYKQ 60

```
Qy 61 NRSYERLALVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 120
Db 61 NRSYERLALVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 120
Qy 121 VMLEPRIHKIAILGLGSSIGTPPEGITAEVLVVTSPFDELQRRASERAGKIIVYNO 175
Db 121 VMLEPRIHKIAILGLGSSIGTPPEGITAEVLVVTSPFDELQRRASERAGKIIVYNO 175

RESULT 2
Qy5X6 PRELIMINARY; PRT; 541 AA.
ID Qy5X6 PRELIMINARY; PRT; 541 AA.
AC Qy5X6;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Blood plasma glutamate carboxypeptidase precursor (EC 3.4.17.21).
GN PGCP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99223495; PubMed=10206990;
RA Gingras R., Richard C., El-Alfy M., Morales C.R., Potier M.,
RA Pehezhetsky A.V.;
RT "Purification, cDNA cloning, and expression of a new human blood
RT plasma glutamate carboxypeptidase homologous to N-acetyl-aspartyl-
RT alpha-glutamate carboxypeptidase/prostate-specific membrane antigen.";
RL J. Biol. Chem. 274:11742-11750(1999).
DR EMBL; AF119386; AAD31418.1; -.
DR MEROPS; M28.014; -.
DR InterPro; IPR003137; PA.
DR PROSITE; PS50840; PA; 1.
KW Carboxypeptidase; Hydrolase; Signal.
FT SIGNAL 1 24
FT CHAIN 45 541 BLOOD PLASMA GLUTAMATE CARBOXYPEPTIDASE.
SQ SEQUENCE 541 AA; 59931 MW; 98138FEE97081F6B CRC64;

Query Match 100.0%; Score 883; DB 4; Length 541;
Best Local Similarity 100.0%; Pred. No. 2.3e-75;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKFLIFAFGGVHLSTLCSGKAICNGISKRTPEEIKKEIASCGVAAIINLAVYGAQ 60
Db 1 MKFLIFAFGGVHLSTLCSGKAICNGISKRTPEEIKKEIASCGVAAIINLAVYGAQ 60
Qy 61 NRSYERLALVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 120
Db 61 NRSYERLALVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 120
Qy 121 VMLEPRIHKIAILGLGSSIGTPPEGITAEVLVVTSPFDELQRRASERAGKIIVYNO 175
Db 121 VMLEPRIHKIAILGLGSSIGTPPEGITAEVLVVTSPFDELQRRASERAGKIIVYNO 175

RESULT 3
Q8NBZ1 PRELIMINARY; PRT; 472 AA.
ID Q8NBZ1 PRELIMINARY; PRT; 472 AA.
AC Q8NBZ1;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical protein FLJ90651.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Placenta;
RA Isegai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
```

```
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto Y., Makamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuho J., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Nimomiya K.;
RT "NDO human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK075132; BAC11423.1; -.
DR InterPro; IPR003137; PA.
DR PROSITE; PS50840; PA; 1.
KW Hypothetical protein.
SQ SEQUENCE 472 AA; 51784 MW; 83F2596F5DC52P3 CRC64;

Query Match 98.8%; Score 872; DB 4; Length 472;
Best Local Similarity 98.9%; Pred. No. 2.1e-74;
Matches 173; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKFLIFAFGGVHLSTLCSGKAICNGISKRTPEEIKKEIASCGVAAIINLAVYGAQ 60
Db 1 MKFLIFAFGGVHLSTLCSGKAICNGISKRTPEEIKKEIASCGVAAIINLAVYGAQ 60
Qy 61 NRSYERLALVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 120
Db 61 NRSYERLALVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 120
Qy 121 VMLEPRIHKIAILGLGSSIGTPPEGITAEVLVVTSPFDELQRRASERAGKIIVYNO 175
Db 121 VMLEPRIHKIAILGLGSSIGTPPEGITAEVLVVTSPFDELQRRASERAGKIIVYNO 175

RESULT 4
Q70216 PRELIMINARY; PRT; 433 AA.
ID Q70216 PRELIMINARY; PRT; 433 AA.
AC Q70216;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hematopoietic lineage switch 2.
GN PGP OR HUS2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Williams J.H., Chan C.-Y., Klinken S.P.;
RT "Hematopoietic lineage switch 2 (HUS2), a novel mRNA species induced
RT during an erythroid to myeloid lineage switch.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF009513; AAC17945.1; -.
DR MEROPS; M28.014; -.
DR MGD; MGI:1889205; PgcP.
DR InterPro; IPR001464; Anmexin.
DR PROSITE; PS00223; ANMEXIN; 1.
DR PROSITE; PS50840; PA; 1.
SQ SEQUENCE 433 AA; 47858 MW; 89C9B93712004669 CRC64;

Query Match 83.4%; Score 736; DB 11; Length 433;
Best Local Similarity 85.7%; Pred. No. 1.1e-61;
Matches 150; Conservative 8; Mismatches 15; Indels 2; Gaps 1;

Qy 1 MKFLIFAFGGVHLSTLCSGKAICNGISKRTPEEIKKEIASCGVAAIINLAVYGAQ 60
Db 1 MRSLFLEFT--VHLALSGKAVKNGVSGRTREIEKEELANVEDVAKAIINLAVYKQ 58
Qy 61 NRSYERLALVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 120
Db 59 NRSYERGLVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHPVRIIPMERGESGA 118
Qy 121 VMLEPRIHKIAILGLGSSIGTPPEGITAEVLVVTSPFDELQRRASERAGKIIVYNO 175
Db 119 VMLEPRIHKIAILGLGSSIGTPPEGITAEVLVVTSPFDELQRRASERAGKIIVYNO 173
```

RESULT 5  
Q9WVJ3 PRELIMINARY; PRT; 470 AA.  
AC Q9WVJ3;  
DT 01-NOV-1999 (Tremblrel. 12, Created)  
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Aminopeptidase (Similar to plasma glutamate carboxypeptidase).  
GN PCP.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Liu C.H., Lin B.Y., Chang L.Y.;  
RT "Cloning of the mouse aminopeptidase gene."  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon;  
RA Strauberg R.;  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Body; and Mesonephros;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium;  
RT the RIKEN Genome Exploration Research Group Phase I & II Team;  
RL "Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs."  
RT Nature 420:563-573(2002).  
DR EMBL: AF107835; AAD43215.1; -  
DR EMBL: BC037067; AAH37067.1; -  
DR EMBL: AK032972; BAC28105.1; -  
DR EMBL: AK075686; BAC35891.1; -  
DR MEROPS: M28.014; -  
DR MGD: MGI:1889205; PgcP.  
DR InterPro: IPR001464; Annexin.  
DR InterPro: IPR003137; PA.  
DR PROSITE: PS00223; ANNEXIN; 1.  
DR PROSITE: PS50840; PA; 1.  
KM Carboxypeptidase.  
SQ SEQUENCE 470 AA; 51813 MW; 0F3490681691866A CRC64;  
Query Match 83.4%; Score 736; DB 11; Length 470;  
Best Local Similarity 85.7%; Pred. No. 1.8e-61;  
Matches 150; Conservative 8; Mismatches 15; Indels 2; Gaps 1;  
QY 1 MKFLIFAFPGVHLLSCGKAIKNGISKRTFEIKETIASCGDVAKAIINLAVYGAQ 60  
DB 1 MRSIFFLFPI--VHLALGSGAVKPNVSVQRTFEIKETIASYEDVAKAIINLAVYGYQ 58  
QY 61 NRSYERLALVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHLEVRIPMERGES 120  
DB 59 NRSYERLGLVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHLEVRIPMERGES 118  
QY 121 VMEPRIHKIALIGLSSIGTPPEGITAEVLVVSFDELQRRASEARAGKIIVNQ 175  
DB 119 VMEPRIHKIALIGLSSIGTPPEGITAEVLVVSFDELQRRASEARAGKIIVNQ 173  
RESULT 6  
Q9JLVO PRELIMINARY; PRT; 472 AA.  
AC Q9JLVO;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Liver annexin-like protein.  
GN LAL.

OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN RP SEQUENCE FROM N.A.  
RA Servillo G., Della Fazio M.A., Piobbico D., Bartoli D., Castellani M.,  
RA Brancorsini S., Viola Magni M.,  
RT "LAL, a novel gene involved during liver regeneration."  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF131077; AAF36518.1; -  
DR MEROPS: M28.014; -  
DR InterPro: IPR001464; Annexin.  
DR InterPro: IPR003137; PA.  
DR PROSITE: PS00223; ANNEXIN; 1.  
DR PROSITE: PS50840; PA; 1.  
SQ SEQUENCE 472 AA; 52011 MW; 55C68FB90E63F265 CRC64;  
Query Match 83.1%; Score 734; DB 11; Length 472;  
Best Local Similarity 85.1%; Pred. No. 2.8e-61;  
Matches 149; Conservative 8; Mismatches 18; Indels 0; Gaps 0;  
QY 1 MKFLIFAFPGVHLLSCGKAIKNGISKRTFEIKETIASCGDVAKAIINLAVYGAQ 60  
DB 1 MRSIFFLFVAVVHLFSLGSGKAIYKSGVSVQRTFEIKETIASYEDVAKAIINLAVYGYQ 60  
QY 61 NRSYERLALVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHLEVRIPMERGES 120  
DB 61 NRSYERLGLVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHLEVRIPMERGES 120  
QY 121 VMEPRIHKIALIGLSSIGTPPEGITAEVLVVSFDELQRRASEARAGKIIVNQ 175  
DB 121 VMEPRIHKIALIGLSSIGTPPEGITAEVLVVSFDELQRRASEARAGKIIVNQ 175  
RESULT 7  
Q9Z1Y1 PRELIMINARY; PRT; 472 AA.  
AC Q9Z1Y1;  
DT 01-MAY-1999 (Tremblrel. 10, Created)  
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Hematopoietic lineage switch 2 related protein.  
GN HLS2-RP.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN RP SEQUENCE FROM N.A.  
RC STRAIN=Fisher;  
RA Chen Y., Talmage D.;  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF097723; AAC72384.1; -  
DR MEROPS: M28.014; -  
DR InterPro: IPR001464; Annexin.  
DR InterPro: IPR003137; PA.  
DR PROSITE: PS00223; ANNEXIN; 1.  
DR PROSITE: PS50840; PA; 1.  
SQ SEQUENCE 472 AA; 51969 MW; DC2111651879E6E2 CRC64;  
Query Match 82.3%; Score 727; DB 11; Length 472;  
Best Local Similarity 84.6%; Pred. No. 1.3e-60;  
Matches 148; Conservative 8; Mismatches 19; Indels 0; Gaps 0;  
QY 1 MKFLIFAFPGVHLLSCGKAIKNGISKRTFEIKETIASCGDVAKAIINLAVYGAQ 60  
DB 1 MRSIFFLFVAVVHLFSLGSGKAIYKSGVSVQRTFEIKETIASYEDVAKAIINLAVYGYQ 60  
QY 61 NRSYERLALVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHLEVRIPMERGES 120  
DB 61 NRSYERLGLVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVLHLEVRIPMERGES 120

Qy 121 VMLERRIKHAIILGSSIGTPPEGITAEVLVVTSPFDELQRRASERARKIVYNNQ 175  
 Db 121 VVVVPRHKLAIILGSSIGTPPEGITAEVLVVTSPFDELQRRASERARKIVYNNQ 175

RESULT 8

ID Q9UNM8 PRELIMINARY; PRT; 140 AA.  
 AC Q9UNM8:  
 DT 01-MAY-2000 (TRENBLREL. 13, Created)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)  
 DE Aminopeptidase (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Liu C.H., Lin B.Y., Chang L.Y.;  
 RT "Cloning of the human aminopeptidase gene."  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF107833; AAD43213.1;  
 DR MEROPS; M28.014;  
 DR InterPro; IPR003137; PA.  
 DR PROSITE; PSS0840; PA; 1.  
 FT NON\_TER 1  
 FT NON\_TER 140  
 SQ SEQUENCE 140 AA; 15472 MW; F34ACEAB33A1AD24 CRC64;

Query Match 55.8%; Score 493; DB 4; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-39; Indels 0; Gaps 0;  
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 LSGSKNLEKAIQIMYQNLQODGLEKYLHPVRIPIHMERGESAVMLERPHIKHAIILGGS 137  
 Db 1 LSGSKNLEKAIQIMYQNLQODGLEKYLHPVRIPIHMERGESAVMLERPHIKHAIILGGS 60  
 Qy 138 SIGTPPEGITAEVLVVTSPFDELQRRASERARKIVYNNQ 175  
 Db 61 SIGTPPEGITAEVLVVTSPFDELQRRASERARKIVYNNQ 98

RESULT 9

ID Q76552 PRELIMINARY; PRT; 493 AA.  
 AC Q76552:  
 DT 01-NOV-1998 (TRENBLREL. 08, Created)  
 DT 01-NOV-1998 (TRENBLREL. 08, Last sequence update)  
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)  
 DE Aminopeptidase ES-62 precursor.  
 OS Acanthocheilichthys viltate (Dipetalonema viteae).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 OC Onchocercidae; Acanthocheilichthys.  
 NCBI\_TaxID=6277;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20055923; PubMed=10589978;  
 RA Harnett W., Houston K.M., Tate E., Garate T., Apfel H., Adam R.,  
 RA Haslam S.M., Panico M., Paxton T., Dell A., Morris H., Brzeski H.;  
 RT "Molecular cloning and demonstration of an aminopeptidase activity in  
 a filarial nematode glycoprotein."  
 RL Mol. Biochem. Parasitol. 104:11-23(1999).  
 DR EMBL; AF077194; AAC28365.1;  
 DR MEROPS; M28.015;  
 DR InterPro; IPR003137; PA.  
 DR PROSITE; PSS0840; PA; 1.  
 KM Signal.  
 FT CHAIN 1 19 POTENTIAL.  
 FT CHAIN 20 493 AMINOPEPTIDASE ES-62.  
 SQ SEQUENCE 493 AA; 54916 MW; C7E9227DD18E1DF4 CRC64;

Query Match 24.6%; Score 217; DB 5; Length 493;  
 Best Local Similarity 32.5%; Pred. No. 4.2e-12;  
 Matches 55; Conservative 32; Mismatches 58; Indels 24; Gaps 5;

Qy 15 LSLCSKAIC-----KNGISKRTPEIKELIASCGDVAKATINLAAYGKAQNSYERL 67  
 Db 14 LTVVLGAADVLPDKTAPKKNYIOETFGKEVAB-----LIQYTKGEVGLAYQWL 62  
 Qy 68 ALIVDTVGPRLSSKNLEKAIQIMYQNLQODGLEKYLHPVRIPIHMERGESAVMLERPR 126  
 Db 63 SKLVDSGGRHNVGSDSLKSIAPLESLLKNDNPKVTHEVPLPHVVRGNDVEMLEPR 122  
 Qy 127 IHKIAIILGSSIGTPPEGITAEVLVVTSPFDELQRRASERARKIVYNNQ 175  
 Db 123 NQRLNVLALG---GSEPSATGSEVTVYDDDV--KDDVRGKIVYVNAQ 166

RESULT 10

ID Q8EBH7 PRELIMINARY; PRT; 414 AA.  
 AC Q8EBH7:  
 DT 01-MAR-2003 (TRENBLREL. 23, Created)  
 DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)  
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)  
 DE Peptidase, M28D family.  
 GN S03539.  
 OS Shewanella oneidensis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;  
 OC Alteromonadaceae; Shewanella.  
 NCBI\_TaxID=70863;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MR-1;  
 RA MEDLINE=22297686; PubMed=12368813;  
 RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,  
 RA Read T.D., Eisen J.A., Seehadi R., Ward N., Methe B., Clayton R.A.,  
 RA Meyer T., Tsaplin A., Scott J., Beaman M., Brinkac L., Daugherty S.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,  
 RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,  
 RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,  
 RA Mueller J., Khouri H., Gill J., Usterback T.R., McDonald L.A.,  
 RA Feldblyum T.V., Smith H.O., Venter J.C., Neilson K.H., Frazer C.M.;  
 RT "Genome sequence of the dissimilatory metal ion-reducing bacterium  
 Shewanella oneidensis."  
 RL Nat. Biotechnol. 20:1118-1123(2002).  
 DR EMBL; AB015790; AAN56530.1;  
 DR TIGR; SO3539;  
 KM Complete proteome.  
 SQ SEQUENCE 414 AA; 44028 MW; C9AEB2A5773DBCB4 CRC64;

Query Match 20.4%; Score 180.5; DB 16; Length 414;  
 Best Local Similarity 38.8%; Pred. No. 9.8e-09;  
 Matches 40; Conservative 17; Mismatches 45; Indels 1; Gaps 1;

Qy 74 VGPRLLSGSKNLEKAIQIMYQNLQODGLEKYLHPVRIPIHMERGESAVMLERPHIKHAIIL 133  
 Db 3 VGPRLLSGSKNLEKAIQIMYQNLQODGLEKYLHPVRIPIHMERGESAVMLERPHIKHAIIL 62  
 Qy 134 GLGSSIGTPPEGITAEVLVVTSPFDELQRRASERARKIVYNNQ 175  
 Db 63 ALGGSVATPVEGIKAKIARFNSLEALQOATPDDVKGIATFDQ 105

RESULT 11

ID Q9A5A8 PRELIMINARY; PRT; 467 AA.  
 AC Q9A5A8:  
 DT 01-JUN-2001 (TRENBLREL. 17, Created)  
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)  
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)  
 DE Aminopeptidase, putative.  
 GN CC2544.  
 OS Caulobacter crescentus.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;  
 OC Caulobacteraceae; Caulobacter.  
 OX NCBI\_TaxId=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=21173699; PubMed=11259647;  
 RA Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA DeBoy R.T., Dodson R.D., Durkin A.S., Gilm M.L., Haft D.H.,  
 RA Kolonay J.F., Sait J., Craven M.B., Khouli H., Smetly J., Berry K.,  
 RA Salzberg S.L., Venter J.C., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 RA "Complete genome sequence of Caulobacter crescentus.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 RL EMBL: AE005922; AAK24515.1; -  
 DR TIGR: CC2544; -  
 DR InterPro: IPR003137; PA.  
 DR PROSITE: PS50840; PA; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 467 AA; 4898 MW; 9984DE2A9D84F00F CRC64;  
 Query Match 18.5%; Score 163.5; DB 16; Length 467;  
 Best Local Similarity 35.5%; Pred. No. 4.8e-07;  
 Matches 38; Conservative 17; Mismatches 51; Indels 1; Gaps 1;  
 QY 70 LVDTVGRSLGSGKNLEKAIQIWMYONLQODGLEKHLPEVPIPHMERGESAVMLEPRRIH 129  
 DB 49 LTTNIGRLVSGSPAMAKAKMSVAKFKALGFTNKKVDEFAKPSVRGEESAEIWAAYAMK 108  
 QY 130 IAILGSSIGTPPEGITAEVLVVTSPDELQRRASEA-RGKIVVYNO 175  
 DB 109 LGAVGLGRVTVTPAGIGIEAVLAKTRADMAAPDGLAKKIVITQ 155  
 RESULT 12  
 Q8PHE8 PRELIMINARY; PRT; 472 AA.  
 AC O8PHE8;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Antimicrobial.  
 GN XAC3309.  
 OS Xanthomonas axonopodis (pv. citri).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 OX NCBI\_TaxId=92829;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=306 / ATCC 13902 / XV 101;  
 RX MEDLINE=22022145; PubMed=12024217;  
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,  
 RA Camarotte G., Camnava F., Cardozo J., Chambergo F., Ciapina L.P.,  
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorty H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,  
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sana J.A.D., Silva C., de Souza R.F.,  
 RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
 RA Sebubal J.C., Kitajima J.P.;  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities.";  
 RL Nature 417:459-463(2002).  
 DR EMBL: AE011976; AAM38152.1; -  
 DR InterPro: IPR003137; PA.

DR PROSITE: PS50840; PA; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 472 AA; 4878 MW; 9956CDF1F49FC68 CRC64;  
 Query Match 13.7%; Score 121; DB 16; Length 472;  
 Best Local Similarity 27.0%; Pred. No. 0.0053;  
 Matches 44; Conservative 20; Mismatches 75; Indels 24; Gaps 5;  
 QY 12 VHLISCSGKAIKNGI---SKRTFEIKERISCGDVAKAIINLVYGAQRYSYERLA 68  
 DB 8 ISALLACSCALAAQTSTIPDSALRTAAQLREQ-ALADTGRVAVQ----- 50  
 QY 69 LVDTVGRSLGSGKNLEKAIQIWMYONLQODGLEKHLPEVPIPHMERGESAVMLEPRRIH 128  
 DB 51 SLTTEVGRPRAGADPRPRAVAMAKAFASIGFQKWTPEPTPFKMERSSHAIVGHAQ 110  
 QY 129 KIALIGSSIGTPPEGITAEVLVVTSPDELQRR-ASEARGT 170  
 DB 111 PLTTTALGSGSPGCTVEG---EIVRFETLALQAAPAGSLAGKI 150  
 RESULT 13  
 Q8P625 PRELIMINARY; PRT; 375 AA.  
 AC Q8P625;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Antimicrobial.  
 GN XCC3157.  
 OS Xanthomonas campestris (pv. campestris).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Xanthomonas.  
 OX NCBI\_TaxId=340;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33913 / NCPPB 528;  
 RX MEDLINE=22022145; PubMed=12024217;  
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,  
 RA Camarotte G., Camnava F., Cardozo J., Chambergo F., Ciapina L.P.,  
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorty H.,  
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
 RA Pereira H.A., Rossi A., Sana J.A.D., Silva C., de Souza R.F.,  
 RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,  
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
 RA Sebubal J.C., Kitajima J.P.;  
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
 RT host specificities.";  
 RL Nature 417:459-463(2002).  
 DR EMBL: AE012431; AAM42427.1; -  
 DR InterPro: IPR003137; PA.  
 DR PROSITE: PS50840; PA; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 375 AA; 38756 MW; C75FAA4914E2EA05 CRC64;  
 Query Match 13.3%; Score 117; DB 16; Length 375;  
 Best Local Similarity 30.1%; Pred. No. 0.0094;  
 Matches 34; Conservative 16; Mismatches 59; Indels 4; Gaps 2;  
 QY 59 AONRSYERLALVDVTPRSLGSGKNLEKAIQIWMYONLQODGLEKHLPEVPIPHMERGES 118  
 DB 41 ADSTGRVAVSLTTEVGRPRAGADPRPRAVAMAKAFKISIGFQKWTPEPTPFKMERSE 100  
 QY 119 SAVMLEPRRIHKAIALIGSSIGTPPEGITAEVLVVTSPDELQRR-ASEARGT 170  
 DB 101 QAAVIGHAQPLHTALGSGPGCTVEG---EIVRFETLALQAAPAGSLAGKI 150

## RESULT 14

Q9WZ59 PRELIMINARY; PRT; 557 AA.  
 ID Q9WZ59  
 AC Q9WZ59;  
 DT 01-NOV-1999 (TRENBLrel. 12, Created)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
 DE Pericillin-binding protein 2.  
 GN TM0590.  
 OS Thermotoga maritima.  
 OC Bacteria; Thermotogae; Thermotogaceae; Thermotoga.  
 OX NCBI\_TaxID=2336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MSB / DSM 3109;  
 RX MEDLINE=99287316; PubMed=10360571;  
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,  
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
 RA McDonald L., Uitterback T.R., Malek J.A., Linher K.D., Garrett M.M.,  
 RA Stewart A.W., Cooten M.D., Pratt M.S., Phillips C.A., Richardson D.,  
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,  
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from  
 RT genome sequence of Thermotoga maritima."  
 RL Nature 399:323-329 (1999).  
 RL EMBL; AE001734; AAC35675.1; -.  
 DR TIGR; TM0590; -.  
 DR InterPro; IPR001092; HLH\_basic.  
 DR InterPro; IPR005311; PRP\_dimer.  
 DR InterPro; IPR001460; Transpeptidase.  
 DR Pfam; PF03717; PRP\_dimer; 1.  
 DR Pfam; PF09005; Transpeptidase; 1.  
 DR PROSITE; PS00038; HLH\_1; 1.  
 DR Complete proteome.  
 SQ SEQUENCE 557 AA; 61900 MW; DCBF394417124A80 CRC64;  
 Query Match 10.6%; Score 93.5; DB 16; Length 557;  
 Best Local Similarity 26.5%; Pred. No. 2.7;  
 Matches 43; Conservative 26; Mismatches 56; Indels 51; Gaps 11;  
 QY 20 GKAIKNGIS-----KRTFERIK-----EIASCGDVAKAIINLA---VYGRK----- 59  
 DB 58 GKVLAKDEVVYVLDPMWLSIDELKKTGLFTPEILSLVKGESIVIDKARADVISKAGRV 117  
 QY 60 ---QNRSTYRLALVDVTGPRISGSKNLEKA---IQIMYQNL---QDDGLEKVLHP--- 107  
 DB 118 VMDYRKRYEPLA-----PHVGYVNADRVAGVGVESVYDFLGTGDKVAVFEPGSA 170  
 QY 108 -----VRIPHWERGESAIVLEPRIRHKTAIIGLSSIGTPPEGI-----TAEVLVVTSPD 157  
 DB 171 ISSEVLKRP- KPGEVDLTITDTRIQAESL- EKIGNPGSVILSDVTRGHIILASFP 228  
 QY 158 E 158  
 DB 229 E 229

## RESULT 15

Q22257 PRELIMINARY; PRT; 1295 AA.  
 ID Q22257  
 AC Q22257;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
 DE T06E4.1 protein.  
 GN T06E4.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Lloyd C.R.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018 (1998).  
 DR EMBL; Z70756; CA94789.1; -.  
 DR WormPep; T06E4.1; CE06368.  
 SQ SEQUENCE 1295 AA; 147395 MW; 57C70DBCCD2172F8 CRC64;

Query Match 10.5%; Score 92.5; DB 5; Length 1295;  
 Best Local Similarity 20.4%; Pred. No. 10;  
 Matches 33; Conservative 44; Mismatches 54; Indels 31; Gaps 5;  
 QY 23 ICKNGISRTPEERKEIASCGDVAKAIINLA VYGRKQNSYERLALVDVTGPRISGSK 82  
 DB 774 VMESVLEKSEFALQLELSASQEVSRSVDAAYOK-----DGLRLVDITLKIKI---E 824  
 QY 83 NLEKAIQIMYQNLQDDGLEKVLHPVRI PHWERGESAIVLEPRI---HKTAIIGLSSI 139  
 DB 825 DTERKA---QDLOQSSVBEIKQLDLQNFQKQVAEVLSEINEXLNSHHRDMDVALASQL 880  
 QY 140 -----GTPPEGITAEVLVVTSPDELORRASEA 166  
 DB 881 EELQHLKLVGESQVENYKEELIGAKINMKEMVDLNAKLGDA 922

Search completed: December 22, 2003, 11:51:54  
 Job time : 28.507 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:49:10 ; Search time 14.0649, Seconds  
(without alignments)  
526.445 Million cell updates/sec

Title: US-09-745-763-36\_COPY\_1\_175  
Perfect score: 883  
Sequence: 1 MKFLIFAFGCVHLLSLCSG.....FDELQRRASEARKIVVYQ 175

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	883	100.0	473	US-09-482-273-120	Sequence 120, App
2	82.5	9.3	337	US-09-107-532A-5998	Sequence 5998, Ap
3	81.5	9.2	262	US-09-328-352-4733	Sequence 4733, Ap
4	81	9.2	812	US-09-198-452A-978	Sequence 978, App
5	79.5	9.0	406	US-09-134-001C-3544	Sequence 3544, Ap
6	78.5	8.9	312	US-08-425-061-18	Sequence 18, Appl
7	78.5	8.9	312	US-08-825-886-18	Sequence 18, Appl
8	78.5	8.9	312	US-08-989-890-18	Sequence 18, Appl
9	78.5	8.9	765	US-08-425-061-19	Sequence 19, Appl
10	78.5	8.9	765	US-08-825-886-19	Sequence 19, Appl
11	78.5	8.9	765	US-08-989-890-19	Sequence 19, Appl
12	78.5	8.9	900	US-08-425-061-20	Sequence 20, Appl
13	78.5	8.9	900	US-08-825-886-20	Sequence 20, Appl
14	78.5	8.9	900	US-08-989-890-20	Sequence 20, Appl
15	78.5	8.9	914	US-08-425-061-21	Sequence 21, Appl
16	78.5	8.9	914	US-08-825-886-21	Sequence 21, Appl
17	78.5	8.9	914	US-08-989-890-21	Sequence 21, Appl
18	78.5	8.9	1202	US-08-425-061-22	Sequence 22, Appl
19	78.5	8.9	1202	US-08-825-886-22	Sequence 22, Appl
20	78.5	8.9	1202	US-08-989-890-22	Sequence 22, Appl
21	78.5	8.9	1363	US-08-425-061-23	Sequence 23, Appl
22	78.5	8.9	1363	US-08-825-886-23	Sequence 23, Appl
23	78.5	8.9	1363	US-08-989-890-23	Sequence 23, Appl
24	78.5	8.9	1852	US-08-425-061-24	Sequence 24, Appl
25	78.5	8.9	1852	US-08-825-886-24	Sequence 24, Appl
26	78.5	8.9	1852	US-08-989-890-24	Sequence 24, Appl
27	78.5	8.9	1863	US-08-425-061-16	Sequence 16, Appl

28	78.5	8.9	1863	1	US-08-598-591-2	Sequence 2, Appl1
29	78.5	8.9	1863	1	US-08-480-784-2	Sequence 2, Appl1
30	78.5	8.9	1863	1	US-08-483-553-2	Sequence 2, Appl1
31	78.5	8.9	1863	1	US-08-487-002-2	Sequence 2, Appl1
32	78.5	8.9	1863	1	US-08-483-554B-2	Sequence 2, Appl1
33	78.5	8.9	1863	1	US-08-798-691-2	Sequence 2, Appl1
34	78.5	8.9	1863	1	US-08-798-691-4	Sequence 2, Appl1
35	78.5	8.9	1863	1	US-08-798-691-6	Sequence 2, Appl1
36	78.5	8.9	1863	1	US-08-488-011B-2	Sequence 2, Appl1
37	78.5	8.9	1863	2	US-08-825-886-16	Sequence 16, Appl1
38	78.5	8.9	1863	2	US-08-603-753D-2	Sequence 2, Appl1
39	78.5	8.9	1863	3	US-08-825-487A-2	Sequence 2, Appl1
40	78.5	8.9	1863	3	US-08-825-487A-4	Sequence 4, Appl1
41	78.5	8.9	1863	3	US-08-825-487A-6	Sequence 6, Appl1
42	78.5	8.9	1863	3	US-09-074-476-2	Sequence 2, Appl1
43	78.5	8.9	1863	3	US-09-074-476-4	Sequence 4, Appl1
44	78.5	8.9	1863	3	US-09-074-476-6	Sequence 6, Appl1
45	78.5	8.9	1863	3	US-09-099-753-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1  
US-09-482-273-120  
Sequence 120, Application US/09482273  
Patent No. 6534631  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 71 Human Secreted Proteins  
FILE REFERENCE: P2030P1  
CURRENT APPLICATION NUMBER: US/09/482,273  
EARLIER FILING DATE: 2000-01-13  
EARLIER APPLICATION NUMBER: PCT/US99/15849  
EARLIER FILING DATE: 1999-07-14  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/092,922  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/092,956  
EARLIER FILING DATE: 1998-07-15  
NUMBER OF SEQ ID NOS: 267  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 120  
LENGTH: 473  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (473)  
OTHER INFORMATION: Xaa equals stop translation  
US-09-482-273-120  
Query Match 100.0%; Score 883; DB 4; Length 473;  
Best Local Similarity 100.0%; Pred. No. 1.1e-93;  
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKFLIFAFGCVHLLSLCSGKAICTKNGISKTPEIRIEBTASGCDVAKAIINLAVYKAQ 60  
DB 1 MKFLIFAFGCVHLLSLCSGKAICTKNGISKTPEIRIEBTASGCDVAKAIINLAVYKAQ 60  
QY 61 NRSYERLALVDVTPGPRLSGSKNLEKAIQIMYQWLODDGKELFVRIIPHWERGESA 120  
DB 61 NRSYERLALVDVTPGPRLSGSKNLEKAIQIMYQWLODDGKELFVRIIPHWERGESA 120  
QY 121 VMLPRLHKLIALILGLSSIGTPPEGITAEVLVYVSPDELQRRASEARKIVVYQ 175  
DB 121 VMLPRLHKLIALILGLSSIGTPPEGITAEVLVYVSPDELQRRASEARKIVVYQ 175  
RESULT 2  
US-09-107-532A-5998  
Sequence 5998, Application US/09107532A

Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Arinello, Pamela Deneka  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5998:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 337 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (8) LOCATION 1...337  
SEQUENCE DESCRIPTION: SEQ ID NO: 5998:  
US-09-107-532A-5998  
Query Match  
Best Local Similarity 9.3%; Score 82.5; DB 4; Length 337;  
Matches 43; Conservative 30; Mismatches 44; Indels 99; Gaps 9;  
QY 7 AFGGVHLISCSGKALCKNGISKTFEIKET-----ASGDAVAKAIIN 52  
DB 2 SLIGGIKNN-----KDLVAKAVYDVVKDITLBEVQELLEPKSEHSDVAFPAFS 53  
QY 53 LA-VYGKQNSRYERLALIVDT-----VGPRLSGSKNLE---KAIQIMYONLQODG 99  
DB 54 LAKVTRKAPQADLALEKIDISANFEKLEVVGPYINFPFNKELISKVQLQTVVKEKEYG 113  
QY 100 -----LEKVALEPRIPH---W--- 113  
DB 114 DSNIGNQGVPIIDMSPNIAKPIISMGHLRSTVIGNSIGRIMKIGVQPIRINHLDWGTQ 173  
QY 114 -----ERGESAVMLEP-----RIHKIA 131  
DB 174 FGKLIVAYKKWGTBEAVVKAEPINELLRLVYQFHEVA 209

RESULT 3  
US-09-328-352-4733  
Sequence 4733, Application US/09328352  
Patent No. 6562958

GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 4733  
LENGTH: 262  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-4733  
Query Match  
Best Local Similarity 9.2%; Score 81.5; DB 4; Length 262;  
Matches 31; Conservative 20; Mismatches 34; Indels 11; Gaps 5;  
QY 67 LALIVDTVGPRLSGSKNLEKAIQIMYONLQODGLEKVALEPRIP---HMERGESAVML 123  
DB 85 LSTIVQTRTP-LDGRFALVALNI-----LQIPQLQPISTLQ-VKMPNDLYSTGKMGGLIV 138  
QY 124 EPRHKIALLGLSSSIGTPPEGITAEVLVVTSPDEL 159  
DB 139 EPLSQHQAIVGVGINLKTTP--VTESDQPISTLEDL 172  
RESULT 4  
US-09-198-452A-978  
Sequence 978, Application US/09198452A  
Patent No. 6559294  
GENERAL INFORMATION:  
APPLICANT: Griffois, R.  
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
FILE REFERENCE: 9710-003-999  
CURRENT APPLICATION NUMBER: US/09/198,452A  
CURRENT FILING DATE: 1998-11-24  
NUMBER OF SEQ ID NOS: 6849  
SEQ ID NO 978  
LENGTH: 812  
TYPE: PRT  
ORGANISM: Chlamydia pneumoniae  
US-09-198-452A-978  
Query Match  
Best Local Similarity 9.2%; Score 81; DB 4; Length 812;  
Matches 48; Conservative 28; Mismatches 62; Indels 60; Gaps 9;  
QY 17 LCGSKALCKNGISKTFEIKETIA-SGDAVAKAIINL-----AVYKAQ 60  
DB 460 VCGGKS-CEHDSILSAQHVSKYISPEFYDVSYFINRQGMRTGKOPPHLIBETQDSP 518  
QY 61 NRSYERLAL-VDTVGPRLSGSKNLEKAIQIMYONLQOD----- 98  
DB 519 LSSSIASLAKVDDLFPVLHGFEGEDGTIOGFELIKPVAQPSLSLAATAMDLLTKRI 578  
QY 99 ---GLEKVALEPRIPHERGESAV-MLE---PRIHKIALLGLSSSIGTPPEGITA 148  
DB 579 ASAVGVVVPVQPIINLTFWKRNPETLQNLIEFSPFMIYKTA--HLGSSIG----- 628  
QY 149 EVLVVTSFDELQRRASEA 166  
DB 629 -IFLVKDELEQEKISA 645  
RESULT 5  
US-09-134-001C-3544  
Sequence 3544, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
;; FILE REFERENCE: GTC-007  
;; CURRENT APPLICATION NUMBER: US/09/134,001C  
;; PRIOR FILING DATE: 1998-08-13  
;; PRIOR APPLICATION NUMBER: US 60/064,964  
;; PRIOR FILING DATE: 1997-11-08  
;; PRIOR APPLICATION NUMBER: US 60/055,779  
;; PRIOR FILING DATE: 1997-08-14  
;; NUMBER OF SEQ ID NOS: 5674  
;; SEQ ID NO 3544  
;; LENGTH: 406  
;; TYPE: PRT  
;; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3544

Query Match 9.0%; Score 79.5; DB 4; Length 406;  
Best Local Similarity 24.0%; Pred. No. 0.97;  
Matches 44; Conservative 34; Mismatches 58; Indels 47; Gaps 9;

QY 3 FLIPAFGVLHLS--LCSGKAIKNGISKRTFEEIKETIASCGDAKAI-----NL 53  
DB 87 FLIDIKTYSHLSSKRLVSGNSV-----VYQLEQALSEHFPEADALFNSGDANL 138  
QY 54 AVYGAQNRS-----YERLALVDVTPGRLSGSKNLEKAIQIMYONLQODGLEKVALP 107  
DB 139 AVFIPIFNQNVIVISDQNHASIID--GKLSGSL-----KVIYHLNDLDES-HLAR 189  
QY 108 VRLPHER-----GEES---AVMLEPRIHKAIIIGLSSIGTPREGITAEVLV 152  
DB 190 HTNPDVQKVIYSDVSTNGTKADINRLVHLKQRYNAIILIIDASHSLNLFEVHADIDI 249  
QY 153 VTS 155  
DB 250 VTS 252

## RESULT 6

US-08-425-061-18  
; Sequence 18, Application US/08425061  
; Patent No. 5622829  
; GENERAL INFORMATION:  
; APPLICANT: KING, Mary-Claire  
; APPLICANT: FRIEDMAN, Lori  
; APPLICANT: OSTERMEYER, Beth  
; APPLICANT: ROWELL, Sarah  
; APPLICANT: LYNCH, Eric  
; APPLICANT: SZABO, Csilla  
; APPLICANT: LEE, Ming  
; TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN  
; TITLE OF INVENTION: CANCER  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/425,061  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, Richard A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO  
; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (415) 494-8700  
;; TELEFAX: (415) 494-8771  
;; TELEX: 910 277299  
;; INFORMATION FOR SEQ ID NO: 18:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 312 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-425-061-18

Query Match 8.9%; Score 78.5; DB 1; Length 312;  
Best Local Similarity 22.8%; Pred. No. 0.85;  
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAPFGVHLISLCSGKA---IKNGISKRT-----FEEIKETIAS--CG--DVAKAI 50  
DB 42 IFCKFCMLKLNQKKPSQCPCLCKNDITKSLQESTFSLVELIKIICAFOLDTGLEY 101  
QY 51 INLAVYGAQNRSEYERL--ALLVDVTPGRLSGSKNL-----EKAQIMYONL-- 95  
DB 102 ANSYNPAKKENNSPEHLKDEVSIIGMGYRNRAKRLQSEPPNSLOETISLSVOLSNLGT 161  
QY 96 QODGLEKVALPEYRIP-HMERGESAVMLEPRIHKAIIIGLSS--IGTPREGITAEVLV 152  
DB 162 VRLIRTKRQIQPKTSVYIELSSDS---EDTVNKAITYGSGQDELQITPGTBEI-- 216  
QY 153 VTSFDELQRRASE 165  
DB 217 --SLDSAKKAACE 227

## RESULT 7

US-08-825-886-18  
; Sequence 18, Application US/08825886  
; Patent No. 5821328  
; GENERAL INFORMATION:  
; APPLICANT: KING, Mary-Claire  
; APPLICANT: FRIEDMAN, Lori  
; APPLICANT: OSTERMEYER, Beth  
; APPLICANT: ROWELL, Sarah  
; APPLICANT: LYNCH, Eric  
; APPLICANT: SZABO, Csilla  
; APPLICANT: LEE, Ming  
; TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN  
; TITLE OF INVENTION: CANCER  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/825,886  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/425,061  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, Richard A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 494-8700

TELEFAX: (415) 494-8771  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 312 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-825-886-18

Query Match 8.9%; Score 78.5; DB 2; Length 312;  
Best Local Similarity 22.8%; Pred. No. 0.85;  
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGVHLISLCSGKA---ICKNGISKRT-----FEIEKEIAS--CG---DVAKAI 50  
DB 42 IFCKFCMLKLNQKQPSQCLKNDITKRSIQESTRFSQVLEELKTIKCAFQDGTGLY 101  
QY 51 INLAVYGAQNRSYERL---ALLVDTVGPRLSGSKL-----EKAIQIMYQNL-- 95  
DB 102 ANSYNFAKKENNSPEHLKDEVSIIQSMGYRNARAKLLQSEBPNPQLQETSLVSQLSNLGT 161  
QY 96 QODGLEKXVHLEPVRI-P-HMERGESAVMLPEPHIKHIAILGLGSS--IGTPPEGITAEVLV 152  
DB 162 VRTIARTKRIOPKTSVYIELGSDSS---EDTVNKATYCSVGDQELLQITPGTRDEI-- 216  
QY 153 VTSFDELQRRASE 165  
DB 217 --SLDSAKKAACE 227

## RESULT 8

US-08-989-890-18

Sequence 18, Application US/08989890  
Patent No. 6512091

## GENERAL INFORMATION:

APPLICANT: KING, Mary-Claire

FRIEDMAN, Lori

OSTERMEYER, Beth

ROMELL, Sarah

LYNCH, Eric

SZABO, Csilla

LEE, Ming

TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESS: FLEHR, HOBBACH, TEST, ALBRITTON &amp; HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/989,890

FILING DATE: 12-Dec-1997

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/825,886

FILING DATE: &lt;Unknown&gt;

APPLICATION NUMBER: 08/425,061

FILING DATE: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, Richard A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 312 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-08-989-890-18

Query Match 8.9%; Score 78.5; DB 4; Length 312;  
Best Local Similarity 22.8%; Pred. No. 0.85;  
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGVHLISLCSGKA---ICKNGISKRT-----FEIEKEIAS--CG---DVAKAI 50  
DB 42 IFCKFCMLKLNQKQPSQCLKNDITKRSIQESTRFSQVLEELKTIKCAFQDGTGLY 101  
QY 51 INLAVYGAQNRSYERL---ALLVDTVGPRLSGSKL-----EKAIQIMYQNL-- 95  
DB 102 ANSYNFAKKENNSPEHLKDEVSIIQSMGYRNARAKLLQSEBPNPQLQETSLVSQLSNLGT 161  
QY 96 QODGLEKXVHLEPVRI-P-HMERGESAVMLPEPHIKHIAILGLGSS--IGTPPEGITAEVLV 152  
DB 162 VRTIARTKRIOPKTSVYIELGSDSS---EDTVNKATYCSVGDQELLQITPGTRDEI-- 216  
QY 153 VTSFDELQRRASE 165  
DB 217 --SLDSAKKAACE 227

## RESULT 9

US-08-425-061-19

Sequence 19, Application US/08425061  
Patent No. 5622829

## GENERAL INFORMATION:

APPLICANT: KING, Mary-Claire

FRIEDMAN, Lori

OSTERMEYER, Beth

ROMELL, Sarah

LYNCH, Eric

SZABO, Csilla

APPLICANT: LEE, Ming

TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESS: FLEHR, HOBBACH, TEST, ALBRITTON &amp; HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/425,061

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, Richard A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 494-8700

TELEFAX: (415) 494-8771

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 765 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-425-061-19

Query Match 8.9%; Score 78.5; DB 1; Length 765;  
Best Local Similarity 22.8%; Pred. No. 3.2;  
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGVHLISLSCGK---ICKNGISKRT-----FEIKKEIAS--CG---DVAKAI 50  
DB 42 IFCKFCMLKLNKGKSGQCLCKNDITKSLQESTRFSGVLELKIICAFOLDTGLEY 101  
QY 51 INLAIVGKAONRSYERL---ALLVDVGPPLSGSKNL-----EKAIQIMYONL-- 95  
DB 102 ANSYNPAKKENNSPEHLKDEVSIIQSWGYNRAKRLQSEPNPSIQETSLSVQLSNLGT 161  
QY 96 QODGLEKVLHPVRIIP-HMERGESAVMLEPRHKKIALIGLSS--IGTPEGITAEVLV 152  
DB 162 VRLRTKQRIQPQTSVYIELSDSS---EDTVNKATYCSVGQDELQITPQGTREI-- 216  
QY 153 VTSFDELORRASE 165  
DB 217 --SIDSAKKAACE 227

## RESULT 10

US-08-825-886-19  
Sequence 19, Application US/08825886  
Patent No. 5823328

## GENERAL INFORMATION:

APPLICANT: KING, Mary-Claire  
APPLICANT: FRIEDMAN, Lori  
APPLICANT: OSTERMEYER, Beth  
APPLICANT: ROWELL, Sarah  
APPLICANT: LYNCH, Eric  
APPLICANT: SZABO, Csilla  
TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN  
TITLE OF INVENTION: CANCER  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/825, 886  
FILING DATE:

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/425, 061  
FILING DATE:

## ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, Richard A  
REGISTRATION NUMBER: 36, 627  
REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:  
LENGTH: 765 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-825-886-19

Query Match 8.9%; Score 78.5; DB 2; Length 765;  
Best Local Similarity 22.8%; Pred. No. 3.2;  
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGVHLISLSCGK---ICKNGISKRT-----FEIKKEIAS--CG---DVAKAI 50  
DB 42 IFCKFCMLKLNKGKSGQCLCKNDITKSLQESTRFSGVLELKIICAFOLDTGLEY 101  
QY 51 INLAIVGKAONRSYERL---ALLVDVGPPLSGSKNL-----EKAIQIMYONL-- 95  
DB 102 ANSYNPAKKENNSPEHLKDEVSIIQSWGYNRAKRLQSEPNPSIQETSLSVQLSNLGT 161  
QY 96 QODGLEKVLHPVRIIP-HMERGESAVMLEPRHKKIALIGLSS--IGTPEGITAEVLV 152  
DB 162 VRLRTKQRIQPQTSVYIELSDSS---EDTVNKATYCSVGQDELQITPQGTREI-- 216  
QY 153 VTSFDELORRASE 165  
DB 217 --SIDSAKKAACE 227

## RESULT 11

US-08-989-890-19  
Sequence 19, Application US/08989890  
Patent No. 6512091

## GENERAL INFORMATION:

APPLICANT: KING, Mary-Claire  
APPLICANT: FRIEDMAN, Lori  
APPLICANT: OSTERMEYER, Beth  
APPLICANT: ROWELL, Sarah  
APPLICANT: LYNCH, Eric  
APPLICANT: SZABO, Csilla  
TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN  
TITLE OF INVENTION: CANCER  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/989, 890  
FILING DATE: 12-Dec-1997  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/825, 886  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 08/425, 061  
FILING DATE: <Unknown>

## ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, Richard A  
REGISTRATION NUMBER: 36, 627  
REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 765 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-08-989-890-19

Query Match 8.9%; Score 78.5; DB 4; Length 765;  
Best Local Similarity 22.8%; Pred. No. 3.2;  
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAPFGVHLISLCSGKA---ICKNGISKRT-----PEEIKERIAS--CG---DVAKAI 50  
DB 42 IFCKFCMLKLNQKKGSPCCPLCKNDITKRSIQESTRFSQVLEIKTICAFQDLTGLEY 101  
QY 51 INLAVYGAQRNSYERL---ALLVDTVGPRLSGSKNL-----EKAIQIMYQNL-- 95  
DB 102 ANSYNPAKKENNSPEHLKDEVSIIQSMGYRNRAKRLQSEBPNPSLOETSLSVOLSNLGT 161  
QY 96 QODGLEKXVLEPVRIP-HMERGESAVMLEPRIHKAIIIGLSS--IGTPEEGITAEVLV 152  
DB 162 VRTLTQRRIOPQKTSVYIELGSDS---EDTVNKATYCSVGDQELLQITPGTRDEI-- 216  
QY 153 VTSFDELQRRASE 165  
DB 217 --SLDSAKKAACE 227

RESULT 12  
US-08-425-061-20  
Sequence 20, Application US/08425061  
Patent No. 5622829  
GENERAL INFORMATION:  
APPLICANT: KING, Mary-Claire  
APPLICANT: FRIEDMAN, Lori  
APPLICANT: OSTERMEYER, Beth  
APPLICANT: ROWELL, Sarah  
APPLICANT: LYNCH, Eric  
APPLICANT: SZABO, Csilla  
APPLICANT: LEE, Ming  
TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN  
TITLE OF INVENTION: CANCER  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/425,061  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 900 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-425-061-20

Query Match 8.9%; Score 78.5; DB 1; Length 900;  
Best Local Similarity 22.8%; Pred. No. 4.1;  
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAPFGVHLISLCSGKA---ICKNGISKRT-----PEEIKERIAS--CG---DVAKAI 50  
DB 42 IFCKFCMLKLNQKKGSPCCPLCKNDITKRSIQESTRFSQVLEIKTICAFQDLTGLEY 101  
QY 51 INLAVYGAQRNSYERL---ALLVDTVGPRLSGSKNL-----EKAIQIMYQNL-- 95  
DB 102 ANSYNPAKKENNSPEHLKDEVSIIQSMGYRNRAKRLQSEBPNPSLOETSLSVOLSNLGT 161  
QY 96 QODGLEKXVLEPVRIP-HMERGESAVMLEPRIHKAIIIGLSS--IGTPEEGITAEVLV 152  
DB 162 VRTLTQRRIOPQKTSVYIELGSDS---EDTVNKATYCSVGDQELLQITPGTRDEI-- 216  
QY 153 VTSFDELQRRASE 165  
DB 217 --SLDSAKKAACE 227

RESULT 13  
US-08-825-886-20  
Sequence 20, Application US/08825886  
Patent No. 5821328  
GENERAL INFORMATION:  
APPLICANT: KING, Mary-Claire  
APPLICANT: FRIEDMAN, Lori  
APPLICANT: OSTERMEYER, Beth  
APPLICANT: ROWELL, Sarah  
APPLICANT: LYNCH, Eric  
APPLICANT: SZABO, Csilla  
APPLICANT: LEE, Ming  
TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN  
TITLE OF INVENTION: CANCER  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/825,886  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/425,061  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 900 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-825-886-20

Query Match 8.9%; Score 78.5; DB 2; Length 900;  
Best Local Similarity 22.8%; Pred. No. 4.1;  
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGVHLISLCSGKA---ICKNGISKRT-----FEEIKERIAS--CG---DVAKAI 50  
DB 42 IFCFCMKILNQKKGSPCCPLCKNDITKSLQSTRFSLQVLEBLKICAFQDLTGLEY 101  
QY 51 INLAVYKKAQNRSYERL---ALVDYVGPRLSGSKNL-----EKAIQMYQNL-- 95  
DB 102 ANSYNFAKKENNSPEHLKDEVSIIQSMGYRNARAKRLQSEPPENSLQETSLSVQSLNIGT 161  
QY 96 QQDGLKLVHLEPVRIIP-HMERGESAVMLPRHKAIALIGLSS--IGTPEGITAEVLV 152  
DB 162 VRTIARTQRIQPKQTSYVIELGSDS---EDTVNKATYCSVGDELQITPQGRDEI-- 216  
QY 153 VTSFDELQRRASE 165  
DB 217 --SLDSAKKAACE 227

## RESULT 14

US-08-989-890-20  
Sequence 20, Application US/08989890  
Patent No. 6512091

## GENERAL INFORMATION:

APPLICANT: KING, Mary-Claire  
FRIEDMAN, Lori  
OSTERMAYER, Beth  
ROMELL, Sarah  
LYNCH, Eric  
SZABO, Csilla  
LEE, Ming  
TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN  
CANCER  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/989,890  
FILING DATE: 12-Dec-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/825,886  
FILING DATE: <Unknown>  
APPLICATION NUMBER: 08/425,061  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 900 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-08-989-890-20

Query Match 8.9%; Score 78.5; DB 4; Length 900;  
Best Local Similarity 22.8%; Pred. No. 4.1;  
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGVHLISLCSGKA---ICKNGISKRT-----FEEIKERIAS--CG---DVAKAI 50  
DB 42 IFCFCMKILNQKKGSPCCPLCKNDITKSLQSTRFSLQVLEBLKICAFQDLTGLEY 101  
QY 51 INLAVYKKAQNRSYERL---ALVDYVGPRLSGSKNL-----EKAIQMYQNL-- 95  
DB 102 ANSYNFAKKENNSPEHLKDEVSIIQSMGYRNARAKRLQSEPPENSLQETSLSVQSLNIGT 161  
QY 96 QQDGLKLVHLEPVRIIP-HMERGESAVMLPRHKAIALIGLSS--IGTPEGITAEVLV 152  
DB 162 VRTIARTQRIQPKQTSYVIELGSDS---EDTVNKATYCSVGDELQITPQGRDEI-- 216  
QY 153 VTSFDELQRRASE 165  
DB 217 --SLDSAKKAACE 227

## RESULT 15

US-08-425-061-21  
Sequence 21, Application US/08425061  
Patent No. 5622829

## GENERAL INFORMATION:

APPLICANT: KING, Mary-Claire  
FRIEDMAN, Lori  
OSTERMAYER, Beth  
ROMELL, Sarah  
LYNCH, Eric  
SZABO, Csilla  
LEE, Ming  
TITLE OF INVENTION: GENETIC MARKERS FOR BREAST AND OVARIAN  
CANCER  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/425,061  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-59563-3/DJB/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 914 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:47:05 ; Search time 14.6059 Seconds  
(without alignments)  
2237.645 Million cell updates/sec

Title: US-09-745-763-36\_COPY\_1\_175  
Perfect score: 883  
Sequence: 1 MKRLIFAFGFGVHLISLCSG.....FDEIQRREARSKIVYVYQ 175

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 696363 seqs, 186758610 residues

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- Published Applications\_AA:\*
- 1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*
  - 2: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB.pep:\*
  - 3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep:\*
  - 4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep:\*
  - 5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep:\*
  - 6: /cgn2\_6/ptodata/2/pubppaa/PCTUS\_PUBCOMB.pep:\*
  - 7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep:\*
  - 8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*
  - 9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*
  - 10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*
  - 11: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep:\*
  - 12: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep:\*
  - 13: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*
  - 14: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*
  - 15: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep:\*
  - 16: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep:\*
  - 17: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep:\*
  - 18: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	883	100.0	472	9 US-09-745-763-36	Sequence 36, Appl
2	883	100.0	473	11 US-09-984-871-120	Sequence 120, App
3	92.5	10.5	325	15 US-10-156-761-11911	Sequence 11911, A
4	81.5	9.2	162	12 US-09-882-227-594	Sequence 594, App
5	81	9.2	340	15 US-10-186-886-2	Sequence 2, Appl
6	79	8.5	1390	15 US-10-087-340-35	Sequence 35, Appl
7	78.5	8.9	1863	10 US-09-734-672-2	Sequence 2, Appl
8	78.5	8.9	1863	10 US-09-734-672-4	Sequence 4, Appl
9	78.5	8.9	1863	10 US-09-734-672-6	Sequence 6, Appl
10	78.5	8.9	1863	11 US-09-982-828-2	Sequence 2, Appl
11	78.5	8.9	1863	11 US-09-982-828-4	Sequence 4, Appl
12	78.5	8.9	1863	11 US-09-982-828-6	Sequence 6, Appl
13	78.5	8.9	1863	15 US-10-022-819-2	Sequence 2, Appl
14	74	8.4	384	9 US-09-925-299-1011	Sequence 1011, Ap
15	74	8.4	384	11 US-09-925-299-1011	Sequence 1011, Ap

16	74	8.4	390	8 US-08-731-566-2	Sequence 2, Appl
17	74	8.4	390	11 US-09-823-187-29	Sequence 29, Appl
18	74	8.4	390	11 US-09-823-187-30	Sequence 30, Appl
19	74	8.4	390	12 US-10-094-944-12	Sequence 12, Appl
20	74	8.4	619	11 US-09-036-614A-1	Sequence 1, Appl
21	74	8.4	619	12 US-10-094-749-1964	Sequence 1964, Ap
22	73.5	8.3	253	12 US-09-769-736-44	Sequence 44, Appl
23	73.5	8.3	143	12 US-10-225-068-182	Sequence 182, App
24	73.5	8.3	459	10 US-09-950-772-4	Sequence 4, Appl
25	73.5	8.3	459	15 US-10-289-360-2	Sequence 2, Appl
26	73.5	8.3	1565	12 US-10-262-794A-59	Sequence 59, Appl
27	73.5	8.3	1565	15 US-10-242-056-59	Sequence 59, Appl
28	73	8.3	848	9 US-09-925-301-1161	Sequence 1161, Ap
29	73	8.3	2353	10 US-09-797-862-33	Sequence 33, Appl
30	72.5	8.2	334	9 US-09-159-469-28	Sequence 28, Appl
31	72.5	8.2	334	9 US-09-798-042-28	Sequence 28, Appl
32	72.5	8.2	530	9 US-09-159-469-73	Sequence 73, Appl
33	72.5	8.2	530	9 US-09-798-042-73	Sequence 73, Appl
34	72.5	8.2	590	9 US-09-159-469-64	Sequence 64, Appl
35	72.5	8.2	590	9 US-09-798-042-64	Sequence 64, Appl
36	72	8.2	478	15 US-10-156-761-13744	Sequence 13744, A
37	72	8.2	1217	12 US-10-132-134-2	Sequence 2, Appl
38	71.5	8.1	223	10 US-09-738-626-6926	Sequence 6926, Ap
39	71.5	8.1	435	10 US-09-866-582-33	Sequence 33, Appl
40	71.5	8.1	519	12 US-10-012-697-1524	Sequence 1524, Ap
41	71.5	8.1	600	9 US-09-975-901-2	Sequence 2, Appl
42	71.5	8.1	1288	11 US-09-919-039-209	Sequence 209, App
43	71	8.0	486	12 US-10-381-779-131	Sequence 131, App
44	71	8.0	619	12 US-10-094-749-2525	Sequence 2525, Ap
45	71	8.0	2066	11 US-09-978-244A-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1  
US-09-745-763-36  
Sequence 36, Application US/09745763  
Patent No. US0020065394A1

GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
McCoy, John M.  
Lavallee, Edward R.  
Collins-Racie, Lisa A.  
Evans, Cheryl  
Werberg, David  
Treacy, Maurice  
Spaulding, Vikki

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM

NUMBER OF SEQUENCES: 219  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/745,763  
FILING DATE: 18-Jun-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Sprunger, Suzanne A.  
REGISTRATION NUMBER: 41,323  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8284  
TELEFAX: (617) 876-5851

```

; INFORMATION FOR SEQ ID NO: 36:
;   SEQUENCE CHARACTERISTICS:
;       LENGTH: 472 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: <Unknown>
;       TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-745-763-36

Query Match          100.0%; Score 883; DB 9; Length 472;
Best Local Similarity 100.0%; Pred. No. 4.3e-91;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKFLFAFGVGHLLSLCSGKAICNGISKRTFEETKEIASCGVAATINLAVYGAQ 60
DB 1 MKFLFAFGVGHLLSLCSGKAICNGISKRTFEETKEIASCGVAATINLAVYGAQ 60
QY 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGES 120
DB 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGES 120
QY 121 VMLEPRHKIALILGSSSIGTPPEGITAEVLVVTSPFDELQRRASARGKIIVY 175
DB 121 VMLEPRHKIALILGSSSIGTPPEGITAEVLVVTSPFDELQRRASARGKIIVY 175

RESULT 2
US-09-984-271-120
; Sequence 120, Application US/09984271
; Publication No. US20030040088A1
; GENERAL INFORMATION:
;   APPLICANT: Rosen et al.
;   TITLE OF INVENTION: 71 Human Secreted Proteins
;   FILE REFERENCE: P2030P1
;   CURRENT APPLICATION NUMBER: US/09/984,271
;   CURRENT FILING DATE: 2001-10-29
;   PRIOR APPLICATION NUMBER: 09/482,273
;   PRIOR FILING DATE: 2000-01-13
;   PRIOR APPLICATION NUMBER: PCT/US99/15849
;   PRIOR FILING DATE: 1999-07-14
;   PRIOR APPLICATION NUMBER: 60/092,921
;   PRIOR FILING DATE: 1998-07-15
;   PRIOR APPLICATION NUMBER: 60/092,922
;   PRIOR FILING DATE: 1998-07-15
;   PRIOR APPLICATION NUMBER: 60/092,956
;   PRIOR FILING DATE: 1998-07-15
;   NUMBER OF SEQ ID NOS: 267
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO: 120
;   LENGTH: 473
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (473)
;   OTHER INFORMATION: Xaa equals stop translation
US-09-984-271-120

Query Match          100.0%; Score 883; DB 11; Length 473;
Best Local Similarity 100.0%; Pred. No. 4.3e-91;
Matches 175; Conservative 0; Mismatches .0; Indels 0; Gaps 0;

QY 1 MKFLFAFGVGHLLSLCSGKAICNGISKRTFEETKEIASCGVAATINLAVYGAQ 60
DB 1 MKFLFAFGVGHLLSLCSGKAICNGISKRTFEETKEIASCGVAATINLAVYGAQ 60
QY 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGES 120
DB 61 NRSYERLALVDVTGPRLSGSKNLEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGES 120
QY 121 VMLEPRHKIALILGSSSIGTPPEGITAEVLVVTSPFDELQRRASARGKIIVY 175
DB 121 VMLEPRHKIALILGSSSIGTPPEGITAEVLVVTSPFDELQRRASARGKIIVY 175
```

```

DB 121 VMLEPRHKIALILGSSSIGTPPEGITAEVLVVTSPFDELQRRASARGKIIVY 175

RESULT 3
US-10-156-761-11911
; Sequence 11911, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;   APPLICANT: OMURA, SATOSHI
;   APPLICANT: IKEDA, HARUO
;   APPLICANT: ISHIKAWA, JUN
;   APPLICANT: HORIKAWA, HIROSHI
;   APPLICANT: SHIBA, TADAYOSHI
;   APPLICANT: SAKAKI, YOSHIYUKI
;   APPLICANT: HATTORI, MASAHIRA
;   TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;   FILE REFERENCE: 249-262
;   CURRENT APPLICATION NUMBER: US/10/156,761
;   CURRENT FILING DATE: 2002-05-29
;   PRIOR APPLICATION NUMBER: JP 2001-204089
;   PRIOR FILING DATE: 2001-05-30
;   PRIOR APPLICATION NUMBER: JP 2001-272697
;   PRIOR FILING DATE: 2001-08-02
;   NUMBER OF SEQ ID NOS: 15109
;   SEQ ID NO: 11911
;   LENGTH: 325
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-11911

Query Match          10.5%; Score 92.5; DB 15; Length 325;
Best Local Similarity 21.5%; Pred. No. 0.077;
Matches 37; Conservative 29; Mismatches 73; Indels 33; Gaps 5;

QY 33 FEETKEIASCGVAATINLAV-----YGAONRSYERLALVDVTGPRLSGSKN 83
DB 92 YDQIVTQLAKMAHARALKIKLPVVVRIPYGGIGAVHSESPALPAHVAAGLKVPSPN 151
QY 84 LEKAIQIMYQNLQODGLEKVLHPVRIPIHMERGESAVMLEPRHKIALILGSSSI---- 139
DB 152 ASDAYVMWQQAIOQSD--PVIFPEPKR-RYWDKGVNVEALPDPILKARVVRGGTDLTAA 209
QY 140 -----GTPPEGITAEVLVVTSPFDELQRRASARGKIIVY 173
DB 210 YGPRVVKYCGEAAABEBSKSLVVDLRSMSPIDFVAQASVEKTRLLVVVH 261

RESULT 4
US-09-882-227-594
; Sequence 594, Application US/09882227
; Publication No. US20030158396A1
; GENERAL INFORMATION:
;   APPLICANT: Kleinhous, Harold
;   APPLICANT: Al-Garawi, Amal
;   APPLICANT: Miller, Charles
;   APPLICANT: Tomb, Jean-Francois
;   APPLICANT: Oocmen, Raymond P.
;   TITLE OF INVENTION: Identification of Polynucleotides
;   TITLE OF INVENTION: Encoding No. US20030158396a1 Helicobacter Polypeptides in the
;   FILE REFERENCE: 06132/047002
;   CURRENT APPLICATION NUMBER: US/09/882,227
;   CURRENT FILING DATE: 2001-06-15
;   PRIOR APPLICATION NUMBER: US 08/902,615
;   PRIOR FILING DATE: 1997-07-29
;   NUMBER OF SEQ ID NOS: 638
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO: 594
;   LENGTH: 162
;   TYPE: PRT
;   ORGANISM: Helicobacter pylori
US-09-882-227-594
```

Query Match 9.2%; Score 81.5; DB 12; Length 162;  
Best Local Similarity 23.5%; Pred. No. 0.49;  
Matches 35; Conservative 31; Mismatches 60; Indels 23; Gaps 6;

Qy 28 ISKTFEIEKEIASCGDVAKAIIINLAVYGAONRSYERLALVDTYGPRLSGSKNLEKA 87  
Db 20 IGSFTKENIK-EVAKC-----FMSLSMFLEOEKNOOEFLQDT-----KENOEBL 66

Qy 88 IQIYQNLQODGLEKHLPEVRIPIHMERGESAVMLBERHIKAI---LGLSSIGTTP 143  
Db 67 I---KNIQISIAKNOELEKISFERWENKIOERVL--PKLKRIVTHKQESITSSINTQL 120

Qy 144 EGITAENVLVTSFDELQRPASERAKTIV 172  
Db 121 ESFKDELDSVFEIQKNTQIAYRLAI 149

RESULT 5  
US-10-186-886-2  
; Sequence 2, Application US/10186886  
; Publication No. US20030119061A1  
; GENERAL INFORMATION:  
; APPLICANT: Navia, Manuel A.  
; APPLICANT: Ala, Paul J.  
; APPLICANT: Griffith, James P.  
; APPLICANT: Ali, Janid A.  
; APPLICANT: Faerman, Carlos H.  
; APPLICANT: Moe, Scott T.  
; APPLICANT: Magee, Andrew S.  
; APPLICANT: Connolly, Patrick R.  
; APPLICANT: Perola, Emanuele  
; TITLE OF INVENTION: STRUCTURE-BASED DRUG DESIGN METHODS FOR  
; TITLE OF INVENTION: IDENTIFYING D-ALA-D-ALA LIGASE INHIBITORS AS ANTIBACTERIAL  
; FILE REFERENCE: 10283-014001  
; CURRENT APPLICATION NUMBER: US/10/186,886  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: US 60/301,676  
; PRIOR FILING DATE: 2001-06-28  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 340  
; TYPE: PRT  
; ORGANISM: Chlamydomophila pneumoniae  
US-10-186-886-2

Query Match 9.2%; Score 81; DB 15; Length 340;  
Best Local Similarity 24.2%; Pred. No. 1.6;  
Matches 48; Conservative 28; Mismatches 62; Indels 60; Gaps 9;

Qy 17 LSCGKAICNGISKTFEIEKEIA-SCGDVAKAIIN-----AVYGAQ 60  
Db 9 VCGGS-CEHDSISLAQHSKYISPEFYDVSFYIINQGLMRGKDPPILEETQDGP 67

Qy 61 NRSYERLALL-VDTYGPRLSGSKNLEKAIQIMYQNLQOD----- 98  
Db 68 LSSEIASLAKVDCIFPLHGPFGDGTIQGFPEILGKPYAGPSISLATATMDKLLTKRI 127

Qy 99 -----GLEKVLHPVRIPIHMERGESAV--MLE-----PRIHKAIILGLSSIGTPEGITTA 148  
Db 128 ASAYGVFPVYPQPLNLCFKWKNPELCLIONLIETSPFPIAVKTA--HLSSSIG----- 177

Qy 149 EVLVVTSFDELQRPASER 166  
Db 178 -IFLVDRKEELQEKISEA 194

RESULT 6  
US-10-097-340-35  
; Sequence 35, Application US/10097340  
; Publication No. US20030087250A1  
; GENERAL INFORMATION:

APPLICANT: John MONAHAN  
APPLICANT: Manjula GANNAPARAPU  
APPLICANT: Sebastian HOERSCH  
APPLICANT: Shubhangi KAMATKAR  
APPLICANT: Steve G. KOVATS  
APPLICANT: Rachel E. MEYERS  
APPLICANT: Michael MORRISSEY  
APPLICANT: Peter OLANDT  
APPLICANT: Ami SEN  
APPLICANT: Peter VEIBY  
APPLICANT: Gordon B. MILLS  
APPLICANT: Robert C. BAST, Jr.  
APPLICANT: Karen LU  
APPLICANT: Rosemarie SCHMANDT  
APPLICANT: Xumei ZHAO  
APPLICANT: Karen GLATT  
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer  
; FILE REFERENCE: MRI-030  
; CURRENT APPLICATION NUMBER: US/10/097,340  
; CURRENT FILING DATE: 2002-03-14  
; PRIOR APPLICATION NUMBER: 60/276,025  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/325,149  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/276,026  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/324,967  
; PRIOR FILING DATE: 2001/09/26  
; PRIOR APPLICATION NUMBER: 60/311,732  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/325,102  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/323,580  
; PRIOR FILING DATE: 2001-09-19  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 1390  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-097-340-35

Query Match 8.9%; Score 79; DB 15; Length 1390;  
Best Local Similarity 28.2%; Pred. No. 21;  
Matches 37; Conservative 26; Mismatches 42; Indels 26; Gaps 9;

Qy 15 LSLCGKAICNGISKTFEIEKEIASCGDVAKAIIN---LAVYGAONRSYERLALLV 71  
Db 851 LSLLE-RVLMKDIYTPPOEVEKTVIRKCLEQA-ALVNYGRLSYAKIEENQK-----A 903

Qy 72 DTVGPRLSGSKNLEKAI--QIMYQNLQODGLEKHLPEVRIPIHMERGEE---SAVMLE 124  
Db 904 ENVRLITPAKLEDITRLAEVLVLEVQON--EEHHA-----PVDGGEAFAMWSDLIVE 957

Qy 125 PRIHKAIILGL 135  
Db 958 ---HAETFLSL 965

RESULT 7  
US-09-734-672-2  
; Sequence 2, Application US/09734672  
; Publication No. US20020183268A1  
; GENERAL INFORMATION:  
APPLICANT: Murphy, Patricia D.  
Allen, Antoinette C.  
Alvares, Christopher P.  
Critz, Brenda S.  
Olson, Sheri J.  
Schelter, Deni J.  
Zeng, Bin  
; TITLE OF INVENTION: Coding Sequences of the Human

BRCA1 Gene  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Ave., N.W.  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/734,672  
FILING DATE: 03-Dec-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/966,436  
FILING DATE: 07-No. US20020183268A1-97  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 12-Feb-96  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5055-02-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1863 amino acids  
TYPE: amino acid  
STRANDEDNESS: No. US20020183268A1 Relevant  
TOPOLOGY: No. US20020183268A1 Relevant  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: BRCA1  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-734-672-2

Query Match 8.9%; Score 78.5; DB 10; Length 1863;  
Best Local Similarity 22.8%; Pred. No. 36;  
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAPFGVHLISLCSGKA---ICKNGISKRT-----FEEIKERIAS--CG---DVAKAI 50  
DB 42 IFCKCMKLNQKKGSGCPCLCNDITKSLQESTRFSQVLELKITICAFQDGTGLV 101  
QY 51 INLAVYGAQNRYSYERL--ALLVDTVGPRISGSKNL-----EKAIQIWMYQNL-- 95  
DB 102 ANSYNPAKKENNSPHLKDVEISIIQSMGYRNRAKRLQSEPNPSLQETSLSVQLSNLGT 161  
QY 96 QODGLEKXVHLEPVRI-P-HMERGESAVMLEPRIKIAILGLGSS--IGTPPEGITAEVLV 152  
DB 162 VRLTRTKRIOPKQTSVYIELGSDS---EDTVNKATYCSVGDQLQITPGTRDEI-- 216  
QY 153 VTSFDELORASE 165  
DB 217 --SLDSAKKACE 227

RESULT 8  
US-09-734-672-4  
Sequence 4, Application US/09734672  
Publication No. US20020183268A1  
GENERAL INFORMATION:  
APPLICANT: Murphy, Patricia D.  
Allen, Antonette C.

Alvares, Christopher P.  
Critz, Brenda S.  
Olson, Sheri J.  
Schelter, Denise B.  
Zeng, Bin  
TITLE OF INVENTION: Coding Sequences of the Human  
BRCA1 Gene  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Ave., N.W.  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/734,672  
FILING DATE: 03-Dec-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/966,436  
FILING DATE: 07-No. US20020183268A1-97  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 12-Feb-96  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5055-02-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1863 amino acids  
TYPE: amino acid  
STRANDEDNESS: No. US20020183268A1 Relevant  
TOPOLOGY: No. US20020183268A1 Relevant  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: BRCA1  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-734-672-4

Query Match 8.9%; Score 78.5; DB 10; Length 1863;  
Best Local Similarity 22.8%; Pred. No. 36;  
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAPFGVHLISLCSGKA---ICKNGISKRT-----FEEIKERIAS--CG---DVAKAI 50  
DB 42 IFCKCMKLNQKKGSGCPCLCNDITKSLQESTRFSQVLELKITICAFQDGTGLV 101  
QY 51 INLAVYGAQNRYSYERL--ALLVDTVGPRISGSKNL-----EKAIQIWMYQNL-- 95  
DB 102 ANSYNPAKKENNSPHLKDVEISIIQSMGYRNRAKRLQSEPNPSLQETSLSVQLSNLGT 161  
QY 96 QODGLEKXVHLEPVRI-P-HMERGESAVMLEPRIKIAILGLGSS--IGTPPEGITAEVLV 152  
DB 162 VRLTRTKRIOPKQTSVYIELGSDS---EDTVNKATYCSVGDQLQITPGTRDEI-- 216  
QY 153 VTSFDELORASE 165  
DB 217 --SLDSAKKACE 227

RESULT 9

US-09-734-672-6  
; Sequence 6, Application US/09734672  
; Publication No. US20020183268A1  
GENERAL INFORMATION:  
APPLICANT: Murphy, Patricia D.  
Allen, Antoinette C.  
Alvares, Christopher P.  
Critz, Brenda S.  
Olson, Sheri J.  
Schelter, Denise B.  
Zeng, Bin  
TITLE OF INVENTION: Coding Sequences of the Human  
BRCA1 Gene  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Ave., N.W.  
CITY: Washington  
STATE: District of Columbia  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/734,672  
FILING DATE: 03-Dec-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/966,436  
FILING DATE: 07-No. US20020183268A1-97  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 12-Feb-96  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5055-02-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1863 amino acids  
TYPE: amino acid  
STRANDEDNESS: NO. US20020183268A1 Relevant  
TOPOLOGY: NO. US20020183268A1 Relevant  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: BRCA1  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-734-672-6  
Query Match 8.9%; Score 78.5; DB 10; Length 1863;  
Best Local Similarity 22.8%; Pred. No. 36;  
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;  
QY 5 IFAPFGVHLISLCSGKA---ICKNGISKRT-----FEIKERIAS--CG--DVAKAI 50  
DB 42 IFCKFCMLKLNQKKGSPQCPKCNNDITKRSLSQESTRFSQVLELLKTIICAFOLDTGLEV 101  
QY 51 INLAVYGAQRNSYERL---ALLVTVGPRLSGSKNL-----EKAQIMWYONL-- 95  
DB 102 ANSNVPAKKENNSPEHLKDEVSIIQSMGYRRAKRLQSEPDNBSLQSTSLSVGLSNIGT 161  
QY 96 QODGLEKHLHPVRIP-HWEGEBSAVMLEPRIKIALIGLGS--IGTPPEGITAEVLV 152  
DB 162 VRLTRTKRRIOPKTSVYIEIGSDSS---EDTVNKATYCSVGDDELLQITPQGRDEL-- 216

QY 153 VTSFDELORRASE 165  
DB 217 --SLDSAKXACE 227  
RESULT 10  
US-09-982-828-2  
; Sequence 2, Application US/09982828  
; Publication No. US20030022184A1  
GENERAL INFORMATION:  
APPLICANT: Murphy, Patricia D.  
Allen, Antoinette C.  
Alvares, Christopher P.  
Critz, Brenda S.  
Olson, Sheri J.  
Thurber, Denise  
Zeng, Bin  
TITLE OF INVENTION: Coding Sequences of the Human  
BRCA1 Gene  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan Lewis & Bockius LLP  
STREET: 1111 Pennsylvania Avenue N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/982,828  
FILING DATE: 22-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/074,453  
FILING DATE: 1998-05-06  
APPLICATION NUMBER: US 08/798,691  
FILING DATE: 1997-02-12  
APPLICATION NUMBER: US 08/598,591  
FILING DATE: 1996-02-12  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael S. Tuscan  
REGISTRATION NUMBER: 43,210  
REFERENCE/DOCKET NUMBER: 44921-5053-01-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-739-3000  
TELEFAX: 202-739-3001  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1863 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: BRCA1 (cm11)  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 17  
MAP POSITION: 17q21  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-982-828-2  
Query Match 8.9%; Score 78.5; DB 11; Length 1863;  
Best Local Similarity 22.8%; Pred. No. 36;  
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;  
QY 5 IFAPFGVHLISLCSGKA---ICKNGISKRT-----FEIKERIAS--CG--DVAKAI 50  
DB 42 IFCKFCMLKLNQKKGSPQCPKCNNDITKRSLSQESTRFSQVLELLKTIICAFOLDTGLEV 101

```

Oy      51  INAVVCKAANRBYERL-----ALIVDVVGRLLSGSKL-----EKAQIOWYOL-- 95
Db      102 ANSNVAFKKKNNSPDEHKODEVSTIGSMGRNPAKRLQSEPPNPBLQETSLSVGLSDJGT 161
Oy      96  QODGLEKXHLPEVRIP-HMERGESAVMLEPPIHKIATILGLGSS--IGTPEPGITAEVLV 152
Db      162 VRTLRKRIQIPQKTSVYIELGSDSS--EDTVNKATVCSVQDQELQITFQGRDEI-- 216
Oy      153 VTFDELDQRRAE 165
Db      217 --SLDSAKKAAE 227

RESULT 11
US-09-982-828-4
/ Sequence 4, Application US/09982828
/ Publication No. US2003022184A1
/ GENERAL INFORMATION:
APPLICANT: Murphy, Patricia D.
          Allen, Antoinette C.
          Alvarez, Christopher P.
          Critz, Brenda S.
          Olsson, Sheri J.
          Thurber, Denise
          Zeng, Bin
TITLE OF INVENTION: Coding Sequences of the Human
                    BRCA1 Gene
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSER: Morgan Lewis & Bockius LLP
STREET: 1111 Pennsylvania Avenue N. W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/982,828
FILING DATE: 22-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/074,453
FILING DATE: 1998-05-06
APPLICATION NUMBER: US 08/798,691
FILING DATE: 1997-02-12
APPLICATION NUMBER: US 08/598,591
FILING DATE: 1996-02-12
ATTORNEY/AGENT INFORMATION:
NAME: Michael S. Tuscan
REGISTRATION NUMBER: 43,210
REFERENCE/DOCKET NUMBER: 44921-5053-01-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-739-3000
TELEFAX: 202-739-3001
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1863 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
STRAIN: BRCA1 (om12)
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 17
MAP POSITION: 17q21
SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

```

US-09-982-828-4
Query Match      8.9%; Score 78.5; DB 11; Length 1663;
Best Local Similarity 22.8%; Pred. No. 36;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAPFGVHILSLSGNA---ICKNGTISKRF-----FEEIKERIAS--CG---DVAKAI 50
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 42 IFCFCFCHKLINQKKGPSQCPCKNDITKSLQSTRFSLQVEULKICAFQDITGLLEY 101
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 51 INLAVYGAQNRSEYERL---ALVDYVGPRLSGSKNL-----EKAQIWMYONL-- 95
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 102 ANSNVPAKKENNSPEHLKDEVSIIIGSNYRNARARLLQSEPNPNSLQETSIVQLSNIGT 161
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 96 QQDLGEKVHLEPVRI P-HMERGESAVMLPEPRIKIALILGSS--IGTPPEGITAEVLY 152
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 162 VRTIRTKRRIQPKQTSVYIEIGSDSS---EDTVAKATVCSVGDELQITPQGTREDEI-- 216
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 153 VTSFDELQRPASE 165
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 217 --SLDSAKKACE 227

RESULT 12
US-09-982-828-6
; Sequence 6, Application US/09982828
; Publication No. US2003022184A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Patricia D.
;           Allen, Antonette C.
;           Alvarez, Christopher P.
;           Critz, Brenda S.
;           Olson, Sheri J.
;           Thurber, Denise
;           Zeng, Bin
; TITLE OF INVENTION: Coding Sequences of the Human
;                   BRCA1 Gene
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan Lewis & Bockius LLP
; STREET: 1111 Pennsylvania Avenue N. W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/982,828
; FILING DATE: 22-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/074,453
; FILING DATE: 1998-05-06
; APPLICATION NUMBER: US 08/798,691
; FILING DATE: 1997-02-12
; APPLICATION NUMBER: US 08/598,591
; FILING DATE: 1996-02-12
; ATTORNEY/AGENT INFORMATION:
; NAME: Michael S. Tuscan
; REGISTRATION NUMBER: 43,210
; REFERENCE/DOCKET NUMBER: 44921-5053-01-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-739-3000
; TELEFAX: 202-739-3001
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1663 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant

```

```

/
/      TOPOLOGY: not relevant
/      MOLECULE TYPE: protein
/      ORIGINAL SOURCE:
/      ORGANISM: Homo sapiens
/      STRAIN: BRCAL (om13)
/      POSITION IN GENOME:
/      CHROMOSOME/SEGMENT: 17
/      MAP POSITION: 17q21
/      SEQUENCE DESCRIPTION: SEQ ID NO: 6:
/
US-09-982-828-6

Query Match      8.9%; Score 78.5; DB 11; Length 1863;
Best Local Similarity 22.8%; Pred. No. 36;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGVHLISLSCGKA---ICKNGISKRT-----FEEIKERIAS--CG---DVAKAI 50
DB 42 IFCKFCMLKILNKKQKSPCKPLCKNDITKRSLOSTFRFSQLVBELKIIICAFOLDTGLEV 101
QY 51 INLAVGAQNRYSYERL---ALLVDTVGPRLSGSKNL-----EKAQIMYONL-- 95
DB 102 ANSYNFAKKENNSPEHLKDEVSIIQSMGYRNRAKRLQSEPNPSLOETSLSVQSLNLGT 161
QY 96 QODGLEKVLHPVRI-P-HMERGESAVMLEPRIHKIALILGLSS--IGTPEGITAEVLV 152
DB 162 VRLTRTKQRIQPKTSVYIELGSDS---EDTVNKATYCSVGQDELIIQITPGTRDEI-- 216
QY 153 VTSFDELQRRASE 165
DB 217 --SLDSAKKACE 227

RESULT 13
US-10-022-819-2
/ Sequence 2, Application US/10022819
/ Publication No. US20030027166A1
/ GENERAL INFORMATION:
/ APPLICANT: ALLEN, Antonette C. P.
/ OLESEN, Sheri J.
/ LAWRENCE, Tammy
/ ANGELLY, Tracy S.
/ RABIN, Mark B.
/ TITLE OF INVENTION: CODING SEQUENCE HAPLOTYPE OF THE HUMAN
/ BRCAL GENE
/ NUMBER OF SEQUENCES: 67
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSEE: Morgan Lewis & Bockius LLP
/ STREET: 1111 Pennsylvania Avenue
/ CITY: Washington DC
/ STATE: District of Columbia
/ COUNTRY: USA
/ ZIP: 20004
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/022,819
/ FILING DATE: 22-Apr-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/074,452
/ FILING DATE: 1998-05-06
/ ATTORNEY/AGENT INFORMATION:
/ NAME: <Unknown>
/ REGISTRATION NUMBER: <Unknown>
/ REFERENCE/DOCKET NUMBER: 044921-5049-01-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-739-3000
/ TELEFAX: 202-739-3001
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:

```

```

/
/      LENGTH: 1863 amino acids
/      TYPE: amino acid
/      STRANDEDNESS: single
/      TOPOLOGY: linear
/      MOLECULE TYPE: protein
/      HYPOTHETICAL: NO
/      ANTI-SENSE: NO
/      FRAGMENT TYPE: N-terminal
/      ORIGINAL SOURCE:
/      ORGANISM: HOMO SAPIENS
/      STRAIN: BRCAL
/      HAPLOTYPE: OM14
/      POSITION IN GENOME:
/      CHROMOSOME/SEGMENT: 17
/      MAP POSITION: 17q21
/      SEQUENCE DESCRIPTION: SEQ ID NO: 2:
/
US-10-022-819-2

Query Match      8.9%; Score 78.5; DB 15; Length 1863;
Best Local Similarity 22.8%; Pred. No. 36;
Matches 44; Conservative 34; Mismatches 76; Indels 39; Gaps 11;

QY 5 IFAFGVHLISLSCGKA---ICKNGISKRT-----FEEIKERIAS--CG---DVAKAI 50
DB 42 IFCKFCMLKILNKKQKSPCKPLCKNDITKRSLOSTFRFSQLVBELKIIICAFOLDTGLEV 101
QY 51 INLAVGAQNRYSYERL---ALLVDTVGPRLSGSKNL-----EKAQIMYONL-- 95
DB 102 ANSYNFAKKENNSPEHLKDEVSIIQSMGYRNRAKRLQSEPNPSLOETSLSVQSLNLGT 161
QY 96 QODGLEKVLHPVRI-P-HMERGESAVMLEPRIHKIALILGLSS--IGTPEGITAEVLV 152
DB 162 VRLTRTKQRIQPKTSVYIELGSDS---EDTVNKATYCSVGQDELIIQITPGTRDEI-- 216
QY 153 VTSFDELQRRASE 165
DB 217 --SLDSAKKACE 227

RESULT 14
US-09-925-299-1011
/ Sequence 1011, Application US/09925299
/ Patent No. US20020055627A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
/ FILE REFERENCE: PA102
/ CURRENT APPLICATION NUMBER: US/09/925,299
/ CURRENT FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/05883
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 1556
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 1011
/ LENGTH: 384
/ TYPE: FRT
/ ORGANISM: Homo sapiens
/
US-09-925-299-1011

Query Match      8.4%; Score 74; DB 9; Length 384;
Best Local Similarity 26.0%; Pred. No. 12;
Matches 34; Conservative 24; Mismatches 43; Indels 30; Gaps 8;

QY 23 ICKNGISKRTFEEIKERIASCG--DVAKAIINLAV---YGR--AQRYSYER-LALLVDVT 73
DB 82 LCGRAL-----EIREKVLGTNNHVDVAKQNNLALLCQNGKKEAVERYYQRLALAYEGQ 135
QY 74 VG---PRLSGSKN-----LEKAIQIMYONLQODGLEKVLHPV-----RIPMERGER 118
DB 136 LGPDNPVAVARTKNNLASCYLKQGYAAEATLYKEILTRAHVQOERGSVDDDHKPIWMAEE 195

```

QY	119	SAVMLEPRHK	129
		:     :	
Db	196	3EEMSKSRHHE	206

```

RESULT 15
US-09-925-299-1011
: Sequence 1011, Application US/09925299
: Publication No. US20030040617A5
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA102.
: CURRENT APPLICATION NUMBER: US/09/925,299
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05883
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 1556
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1011
: LENGTH: 384
: TYPE: prt
: ORGANISM: Homo sapiens
US-09-925-299-1011

```

Query Match	8.4%	Score 74	DB 11	Length 384
Best Local Similarity	26.0%	Pred. No. 12		
Matches 34	Conservative 24	Mismatches 43	Indels 30	Gaps 8

[illegible]

Search completed: December 22, 2003, 11:54:06  
Job time : 15.6059 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:50:05 ; Search time 41 Seconds  
(without alignments)  
38.714 Million cell updates/sec

Title: US-09-745-763-36\_COPY\_231\_240  
Perfect score: 10  
Sequence: 1 TACITVEDAE 10

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100.0	178	24	ABR48107	Human secreted pro
2	100.0	178	24	ABR00275	Human gene 142 enc
3	100.0	472	18	AAW33604	Human secreted pro
4	100.0	472	19	AAW85456	Secreted protein e
5	100.0	472	21	AAH10229	Human fetal kidney
6	100.0	472	21	AAH7254	Human signal pepti
7	100.0	472	21	AAH58879	Human peptidase NA
8	100.0	472	22	AAW93559	Human polypeptide,
9	100.0	472	22	AAE06058	Human gene 18 enco

10	100.0	472	23	ABP61800	Human polypeptide
11	100.0	472	23	ABG33880	Human secreted pro
12	100.0	472	24	ABR47892	Human secreted pro
13	100.0	472	24	ABR48106	Human secreted pro
14	100.0	472	24	ABR00152	Human gene 142 enc
15	100.0	472	24	ABR00274	Human secreted pro
16	100.0	473	21	AAH7081	Human secreted pro
17	100.0	474	22	ABG09405	Novel human diago
18	100.0	481	22	AAH25762	Human protein sequ
19	60.0	328	22	AAH81198	Human RNA-binding
20	60.0	321	22	ABG22827	Novel human diago
21	60.0	323	22	AAH93971	Human stomach cano
22	60.0	333	22	AAH93767	Human protein sequ
23	60.0	335	21	AAH95335	Human quaking spli
24	60.0	328	23	ABP66344	Bifidobacterium 10
25	60.0	341	21	AAH95336	Human quaking spli
26	60.0	344	21	AAH06629	Arabidopsis thalia
27	60.0	344	21	AAH48593	Arabidopsis thalia
28	60.0	377	22	AAH62340	Arabidopsis thalia
29	60.0	417	21	AAH07093	Propionibacterium
30	60.0	417	21	AAH49104	Arabidopsis thalia
31	60.0	424	21	AAH07092	Arabidopsis thalia
32	60.0	424	21	AAH49103	Arabidopsis thalia
33	60.0	425	21	AAH77270	Streptomyces pence
34	60.0	425	21	AAH78822	Amino acid sequenc
35	60.0	432	22	AAH36107	Klebsiella pneumon
36	60.0	433	21	AAH19094	Protein encoded by
37	60.0	448	21	AAH07091	Arabidopsis thalia
38	60.0	448	21	AAH49102	Arabidopsis thalia
39	60.0	532	20	AAH16036	Neurospora crassa
40	60.0	610	22	AAH34063	Staphylococcus aur
41	60.0	618	22	AAH36974	Staphylococcus aur
42	60.0	618	22	AAH37377	Staphylococcus aur
43	60.0	671	23	ABH49703	Listeria monocytop
44	60.0	1068	20	AAH30048	Pancreatic eukaryo
45	60.0	1108	20	AAH30046	Pancreatic eukaryo

#### ALIGNMENTS

RESULT 1	ABR48107	standard; Protein; 178 AA.
ID	ABR48107	
XX	ABR48107	
AC	ABR48107	
XX	12-JUN-2003	(first entry)
DT	12-JUN-2003	
XX	Human secreted protein, SEQ ID 998.	
XX	Cardiant; antiarrhythmic; antiarteriosclerotic; vasotrophic; cyostatic;	
XX	KW vulnerable; antiinflammatory; nootropic; neuroprotective;	
XX	KM antiparkinsonian; gene therapy; human; cardiovascular disorder.	
XX		
OS	Homo sapiens.	
XX		
XX	W0200295010-A2.	
XX		
XX	28-NOV-2002.	
PD	28-NOV-2002.	
XX		
XX	19-MAR-2002; 2002W0-US09785.	
PF	19-MAR-2002; 2002W0-US09785.	
XX		
XX	21-MAR-2001; 2001US-277340P.	
PR	21-MAR-2001; 2001US-277340P.	
XX	19-JUL-2001; 2001US-306171P.	
PR	19-JUL-2001; 2001US-306171P.	
XX	13-NOV-2001; 2001US-331287P.	
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
XX	Human fetal kidney	
XX	Human signal pepti	
PI	Rosen CA, Ruben SM;	
XX		
DR	WPI; 2003-129429/12.	
XX		

PT Novel human secreted proteins, useful for detecting, preventing,  
PT diagnosing, prognosticating, treating and/or ameliorating  
PT cardiovascular disorders such as arrhythmia -  
XX  
XX Claim 13; SEQ ID 998; 1881pp; English.  
XX  
CC The present invention relates to novel human secreted proteins  
CC (ABR47733-ABR48145) and their coding sequences (ACC50344-ACC50856). The  
CC proteins and their coding sequences are useful for the preparation of a  
CC diagnostic or pharmaceutical composition for diagnosing or treating a  
CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,  
CC coronary arteriosclerosis and myocardial ischemia), neural disorders,  
CC immune system disorders, muscular disorders, reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, renal disorders,  
CC proliferative disorders and/or cancerous diseases and conditions, for  
CC wound healing and epithelial cell proliferation, to treat inflammation or  
CC infection, for treating thrombosis and arteriosclerosis, for treating or  
CC preventing neural damage which occurs in neuronal disorders or  
CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
CC disease, to enhance bone and periodontal regeneration and aid in tissue  
CC transplants or bone grafts, to prevent skin aging or hair loss, to  
CC stimulate growth and differentiation of haematopoietic cells and bone  
CC marrow cells when used in combination with other cytokines, to maintain  
CC organs before transplantation or for supporting cell culture of primary  
CC tissues, to increase or decrease differentiation or proliferation of  
CC embryonic stem cells, or to modulate mammalian characteristics or  
CC metabolism.  
CC Note: The sequence data for this patent was published in electronic  
CC format and is available from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
CC  
XX  
SQ Sequence 178 AA;  
Query Match 100.0%; Score 10; DB 24; Length 178;  
Best Local Similarity 100.0%; Pred. No. 0.0023;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 TACTIVEDAE 10  
Db 21 TACTIVEDAE 30  
RESULT 2  
ABR00275  
ID ABR00275 standard; Protein: 178 AA.  
XX  
AC ABR00275;  
XX  
DT 03-APR-2003 (first entry)  
XX  
DE Human gene 142 encoded secreted protein HRACJ35, SEQ ID NO:564.  
XX  
KW Human; secreted protein; digestive disorder; gastrointestinal disorder;  
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;  
KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;  
KW immune disorder; inflammation; infection; wound healing; drug screening;  
KW chromosome identification; chromosome mapping; cytostatic;  
KW antiinflammatory; immunosuppressive; vulnerability; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200276488-A1.  
XX  
PD 03-OCT-2002.  
XX  
PF 19-MAR-2002; 2002WO-US08276.  
XX  
PR 21-MAR-2001; 2001US-277340P.  
PR 19-JUL-2001; 2001US-306171P.  
PR 13-NOV-2001; 2001US-33187P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX

PI Rosen CA, Ruben SM;  
XX  
XX MPI; 2003-029900/02.  
DR N-PSDB; AB271454.  
XX  
XX New human secreted proteins and nucleic acids, useful for detecting,  
PT preventing, diagnosing, prognosticating, treating and/or ameliorating  
PT e.g. gastrointestinal diseases and disorders, or cancers -  
XX  
XX Claim 13; Page 1110; 1216pp; English.  
XX  
CC AB271190-AB271478 represent cDNAs corresponding to 178 human secreted  
CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.  
CC AB271479-AB271540 represent human secreted protein genomic fragments. The  
CC invention also encompasses antibodies specific for the secreted proteins,  
CC the use of the secreted proteins in drug screening, and recombinant  
CC vectors and host cells comprising a nucleic acid of the invention. The  
CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
CC fragments specific for the secreted proteins, and modulators of protein  
CC activity are useful for diagnosing, treating, ameliorating or preventing  
CC digestive disorders. Such conditions include disorders of the mouth,  
CC oesophagus, stomach, small intestine, large intestine, liver, biliary  
CC tract and pancreas, and include cancers of these organs and tissues. The  
CC secreted proteins and their nucleic acids may also be used in the  
CC treatment of immune disorders, inflammation, infection,  
CC hyperproliferative disorders, and to promote wound healing. Nucleic acids  
CC of the invention may be used for chromosome identification, chromosome  
CC mapping, in gene therapy, for identifying individuals from minute  
CC biological samples, as hybridisation probes, and as molecular weight  
CC markers. The present sequence represents a human secreted protein of the  
XX invention.  
XX  
SQ Sequence 178 AA;  
Query Match 100.0%; Score 10; DB 24; Length 178;  
Best Local Similarity 100.0%; Pred. No. 0.0023;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 TACTIVEDAE 10  
Db 21 TACTIVEDAE 30  
RESULT 3  
AAM33604  
ID AAM33604 standard; Protein: 472 AA.  
XX  
AC AAM33604;  
XX  
DT 21-MAY-1998 (first entry)  
XX  
DE Human secreted protein AM282 full-length sequence.  
XX  
KW Secreted protein; AM282; cytokine; human.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FH Peptide 1..24  
FT /label= Sig\_peptide  
FT 25..472  
FT Protein /label= Mat\_protein  
XX  
XX  
XX WO9739030-A2.  
XX  
XX 23-OCT-1997.  
XX  
XX 16-APR-1997; 97WO-US06475.  
XX  
XX 13-JAN-1997; 97US-0783520.  
XX  
XX 18-APR-1996; 96US-0634325.  
XX  
XX (GEMY ) GENETICS INST INC.  
XX

XX Jacobs K, Lavallie ER, McCoy JM, Werberg D, Racie LA;  
 PI Spaulding V;  
 XX WPI; 1997-526400/48.  
 DR N-PSDB; AAV02296.  
 XX  
 PT New isolated secretory proteins AM340, AM282 and AK583 - possibly  
 PT have cytokine, cell proliferation/differentiation regulating,  
 PT immunomodulating activities, etc.  
 XX  
 PS Claim 16; Page 45-47; 59pp; English.  
 XX  
 CC This human secreted protein, designated AM282, is encoded by a  
 CC full-length cDNA clone (see AAV02296), deposited in ATCC 98026, that  
 CC was identified from a database search using an isolated partial  
 CC AM282 clone (see AAT97398). AM282 protein can be used in a claimed  
 CC method for preventing, treating or ameliorating a medical  
 CC condition. It may exhibit cytokine, cell proliferation (either  
 CC inducing or inhibiting) or cell differentiation (either inducing or  
 CC inhibiting) activity or may induce production of other cytokines in  
 CC certain cell populations. It may also exhibit e.g. immune  
 CC stimulating or suppressing activity, haematopoiesis regulating  
 CC activity, tissue growth activity, activin/inhibin activity,  
 CC chemotactic or chemokinetic activity, haemostatic or thrombolytic  
 CC activity, receptor/ligand activity, anti-inflammatory activity,  
 CC tumour inhibition activity, or other activities. No evidence of  
 CC any of these activities is given in the specification.  
 XX  
 SQ Sequence 472 AA;  
 XX  
 Query Match 100.0%; Score 10; DB 18; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TACTIVEDAE 10  
 DB 231 TACTIVEDAE 240  
 XX  
 RESULT 4  
 AAW85456 standard; Protein; 472 AA.  
 ID AAW85456  
 AC AAW85456;  
 XX  
 DT 25-FEB-1999 (first entry)  
 XX  
 DE Secreted protein encoded by clone bu45\_2.  
 XX  
 KM Secreted protein; nutritional activity; immune stimulating; vaccine;  
 KM suppressing activity; haematopoiesis regulating activity;  
 KM tissue growth activity; activin; inhibin activity; chemotaxis;  
 KM chemokinetic activity; haemostasis; thrombolytic activity; receptor;  
 KM ligand; anti-inflammatory; cadherin; tumour invasion suppressor;  
 KM tumour inhibition; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9842739-A2.  
 XX  
 PD 01-OCT-1998.  
 XX  
 PF 20-MAR-1998; 98WO-US05653.  
 XX  
 PR 19-MAR-1998; 98US-0044466.  
 PR 21-MAR-1997; 97US-0822167.  
 XX  
 PA (GEMV ) GENETICS INST INC.  
 PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Werberg D;  
 PI Racie LA, Spaulding V, Treacy M,  
 XX

DR WPI; 1998-609890/51.  
 DR N-PSDB; AAV82779.  
 XX  
 PT New polynucleotides encoding secreted human proteins - derived from  
 PT human foetal brain, adult brain, foetal kidney, placenta or adult  
 PT pineal gland cDNA libraries.  
 XX  
 PS Claim 14; Page 70-72; 113pp; English.  
 XX  
 CC The present sequence represents a secreted protein. The polynucleotide  
 CC and secreted protein are predicted to have biological activities which  
 CC would make them suitable for treating, preventing or ameliorating medical  
 CC conditions in humans and animals, although no supporting data is given.  
 CC Suggested activities include nutritional activity, immune stimulating  
 CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating  
 CC activity, tissue growth activity, activin/inhibin activity,  
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,  
 CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour  
 CC invasion suppressor activity, and tumour inhibition activity (no data is  
 CC given in the specification to support these activities). The  
 CC polynucleotide is also stated to be useful for gene therapy.  
 XX  
 SQ Sequence 472 AA;  
 XX  
 Query Match 100.0%; Score 10; DB 19; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TACTIVEDAE 10  
 DB 231 TACTIVEDAE 240  
 XX  
 RESULT 5  
 AAB10229 standard; Protein; 472 AA.  
 ID AAB10229  
 AC AAB10229;  
 XX  
 DT 16-NOV-2000 (first entry)  
 XX  
 DE Human fetal kidney protein fragment AM282\_11.  
 XX  
 KM Secreted protein; cytotactic; immunostimulatory; antimicrobial;  
 KM antiviral; immunosuppressive; antiinflammatory; vulnerrary; cytokine;  
 KM cell proliferation; differentiation; regulator; treatment; tumor;  
 KM autoimmune disease; inflammatory disorder; wound; microbial infection;  
 KM viral disease; graft versus host reaction suppression.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200037630-A1.  
 XX  
 PD 29-JUN-2000.  
 XX  
 PF 22-DEC-1999; 99WO-US31005.  
 XX  
 PR 23-DEC-1998; 98US-0220876.  
 XX  
 PA (GEMV ) GENETICS INST INC.  
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Werberg D, Treacy M, Bowman MR;  
 XX  
 DR WPI; 2000-442661/38.  
 DR N-PSDB; AAA40493.  
 XX  
 PT Secreted human proteins AS296-11 and AS34-11, useful for treating  
 PT tumors, autoimmune diseases, inflammatory disorders, wounds, microbial  
 PT infections and viral diseases -  
 XX  
 PS Disclosure; Page 199-200; 293pp; English.  
 XX

CC This invention describes novel secreted human proteins (I) which have  
 CC cytosolic, immunostimulatory, antimicrobial, antiviral,  
 CC immunosuppressive, antiinflammatory and vulnereary activity and which act  
 CC as cytokine, cell proliferation or differentiation regulators. (I)  
 CC is useful for treating tumors, autoimmune diseases, inflammatory  
 CC disorders, wounds, microbial infections and viral diseases. (I) is also  
 CC useful for suppressing graft versus host reaction. AAB10226-B10288  
 CC represent the secreted proteins encoded by AAA40490-AA0580 which are  
 CC described in the method of the invention.

XX Sequence 472 AA;  
 SQ

Query Match 100.0%; Score 10; DB 21; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TACTIVEDAE 10  
 |||||  
 DB 231 TACTIVEDAE 240

RESULT 6  
 AAY87254  
 ID AAY87254 standard; Protein; 472 AA.  
 AC AAY87254;  
 XX  
 DT 11-MAY-2000 (first entry)  
 DE Human signal peptide containing protein HSP-31 SEQ ID NO:31.  
 XX  
 KW Human; signal peptide-containing protein; HSP; diagnosis; cancer;  
 KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;  
 KW antimicrobial; neuroprotective; cardiovascular; hepatocytic;  
 KW antiasthmatic; gene therapy; cell proliferation; neurological disorder;  
 KW reproductive disorder; developmental disorder; arteriosclerosis;  
 KW cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;  
 KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;  
 KW Parkinson's disease; Huntington's disease; ovulatory defect;  
 KW muscular dystrophy.

XX Homo sapiens.  
 OS  
 XX  
 PN WO200000610-A2.  
 XX  
 PD 06-JAN-2000.  
 XX  
 XX 25-JUN-1999; 99WO-US14484.  
 XX  
 XX 26-JUN-1998; 98US-0090762.  
 PR 31-JUL-1998; 98US-0094983.  
 PR 01-OCT-1998; 98US-0102686.  
 PR 11-DEC-1998; 98US-0112129.  
 XX  
 PA (INCY-) INCYTE PHARM INC.  
 XX  
 PI Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR,  
 PI Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL,  
 PI Bandman O;  
 XX  
 XX WPI; 2000-160673/14.  
 DR N-PSDB; AA298139.  
 XX  
 XX New human signal peptide-containing proteins useful in treatment,  
 PT prevention and diagnosis of e.g. cancer, inflammation and  
 PT cardiovascular disease  
 XX  
 PS Claim 1; Page 179-180; 327pp; English.  
 XX  
 CC AA298109 to AA298242 encode AAY87224 to AAY87357 which represent the  
 CC human signal peptide-containing proteins HSP-1 to HSP-134. HSPs have  
 CC anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatocytic,  
 CC neuroprotective, cardiovascular and antiasthmatic activities, and can

CC be used in gene therapy. HSPs can be used to treat or prevent disorders  
 CC associated with decreased activity or function of HSP. Antagonists of  
 CC HSP are used to treat or prevent disorders associated with increased  
 CC activity or function of HSP. Such diseases include cell proliferation  
 CC (including cancer), inflammation, cardiovascular, neurological,  
 CC reproductive or developmental disorders, (e.g. arteriosclerosis,  
 CC cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,  
 CC asthma, Crohn's disease, microbial or other infections, congestive or  
 CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's  
 CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSP  
 CC nucleic acids can be used for the recombinant production of HSP, for  
 CC detecting HSP in standard hybridisation and amplification assays (for  
 CC diagnosis and monitoring), in gene therapy, as antisense,  
 CC triplex-forming or ribozyme therapeutics, for detecting related sequences  
 CC or genetic variations, and for chromosomal mapping. HSP are also used to  
 CC raise specific antibodies (Ab) and to screen for agonists and  
 CC antagonists (potential therapeutic agents). Ab are used to diagnose, or  
 CC monitor, HSP-related diseases (in usual immunoassays), as therapeutic  
 CC antagonists, in competitive drug screens, and for purification of HSP  
 CC from natural sources.

XX Sequence 472 AA;  
 SQ

Query Match 100.0%; Score 10; DB 21; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TACTIVEDAE 10  
 |||||  
 DB 231 TACTIVEDAE 240

RESULT 7  
 AAY58879  
 ID AAY58879 standard; Protein; 472 AA.  
 AC AAY58879;  
 XX  
 DT 08-MAY-2000 (first entry)  
 DE Human peptidase NAALAD-ase IV.  
 XX  
 XX NAALAD-ase IV; N-acetylated alpha-linked acidic dipeptidase; human,  
 KW chromosome 8q21.3; prostate cancer; neurodegenerative disease;  
 KW Alzheimer's disease; schizophrenia; ALS; Parkinson's disease;  
 KW peripheral neuropathy; Huntington's disease; acute brain injury;  
 KW multiple sclerosis; peripheral nerve trauma; ischaemia; dementia;  
 KW therapy; diagnosis; neurotropic; neuroprotective; neuroleptic;  
 KW antiparkinsonian; anticonvulsant; vasotrophic.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 3..24  
 FT Modified-site 121  
 FT /note= "membrane-spanning domain"  
 FT Modified-site 179  
 FT /note= "N-glycosylated"  
 FT Modified-site 353  
 FT /note= "N-glycosylated"  
 FT Modified-site 356  
 FT /note= "N-glycosylated"  
 FT Modified-site 396  
 FT /note= "N-glycosylated"  
 XX  
 PN WO200004157-A2.  
 XX  
 PD 27-JAN-2000.  
 XX  
 XX 14-JUL-1999; 99WO-GB02241.  
 PF 14-JUL-1999; 98GB-0015284.  
 XX  
 PR 14-JUL-1998; 98GB-0015284.

XX (JANC) JANSSEN PHARM NV.  
 XX  
 XX Pangalos M, Neefs JEFM, Peeters DCG;  
 PI  
 XX WPI, 2000-182424/16.  
 DR  
 XX N-PSDB; AAZ58313.  
 XX  
 PT New human N-acetylated alpha-linked acidic dipeptidases for treating  
 PT neural disorders e.g. Alzheimer's disease, schizophrenia and  
 PT Parkinson's disease -  
 PT  
 PS Claim 15; Fig 5; 95bp; English.  
 XX  
 XX The present sequence is that predicted for human N-acetylated  
 CC alpha-linked acidic dipeptidase IV (NALAD-ase IV) on the basis of  
 CC isolated cDNA (see AAZ58313). NALAD-ase IV is predicted to be a  
 CC type II integral membrane protein of mol.wt. 51.9 kDa and pI 5.99.  
 CC Expression was low in all tissues examined by RT-PCR. The  
 CC invention provides human NALAD-ase I, II and IV cDNAs and encoded  
 CC polypeptides, as well as vectors, host cells, transgenic organisms,  
 CC antisense nucleic acids, agonists and antagonists. These are useful  
 CC for treating neural disorders such as Alzheimer's disease,  
 CC schizophrenia, ALS, Parkinson's disease, peripheral neuropathy,  
 CC Huntington's disease, acute brain injury, multiple sclerosis,  
 CC exposure to neurotoxins, peripheral nerve trauma, ischaemia or  
 CC dementia (claimed).  
 CC  
 XX  
 SQ Sequence 472 AA;  
 Query Match 100.0%; Score 10; DB 21; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TACTTVEDAE 10  
 DB 231 TACTTVEDAE 240

RESULT 8  
 AAM93559  
 ID AAM93559 standard; Protein; 472 AA.  
 XX  
 XX AAM93559;  
 AC  
 XX  
 XX 06-NOV-2001 (first entry)  
 DT  
 XX  
 XX Human polypeptide, SEQ ID NO: 3329.  
 DE  
 XX  
 XX Human, full length cDNA; cDNA synthesis; oligo-capping.  
 KM  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX EP1130094-A2.  
 PN  
 XX  
 XX 05-SEP-2001.  
 PD  
 XX  
 XX 07-JUL-2000; 2000EP-0114089.  
 PF  
 XX  
 XX 08-JUL-1999; 99JP-0194486.  
 PR  
 XX  
 XX 11-JAN-2000; 2000JP-0118774.  
 PR  
 XX  
 XX 02-MAY-2000; 2000JP-0183765.  
 PR  
 XX  
 XX (HELI-) HELIX RES INST.  
 PA  
 XX  
 XX Oca T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
 PI  
 XX WPI; 2001-524255/58.  
 DR  
 XX  
 XX N-PSDB; AAK94491.  
 DR  
 XX  
 PT 830 Primers useful for synthesizing full length cDNA clones and their  
 PT use in genetic manipulation -

XX  
 PS Claim 8; SEQ ID NO 3329; 1380bp + sequence listing; English.  
 XX  
 XX The invention relates to primers for synthesizing full length cDNA  
 CC clones. 830 cDNA molecules encoding a human protein have been  
 CC isolated and nucleotide sequences of 5' - and 3' -ends of the cDNA  
 CC molecules have been determined. Primers for synthesizing the full length  
 CC cDNA are useful for clarifying the function of the protein encoded by  
 CC the cDNA. The full length clones were obtained by construction of full  
 CC length enriched cDNA libraries that were synthesised by the oligo-capping  
 CC method. The primers enable the production of the full length cDNA easily  
 CC without any special methods. The present sequence is a polypeptide  
 CC encoded by a full length human cDNA of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in CD-ROM format directly from EPO.  
 CC  
 XX  
 SQ Sequence 472 AA;  
 Query Match 100.0%; Score 10; DB 22; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TACTTVEDAE 10  
 DB 231 TACTTVEDAE 240

RESULT 9  
 AAE06058  
 ID AAE06058 standard; Protein; 472 AA.  
 XX  
 XX AAE06058;  
 AC  
 XX  
 XX 24-SEP-2001 (first entry)  
 DT  
 XX  
 XX Human gene 18 encoded secreted protein HRA0335, SEQ ID NO:120.  
 DE  
 XX  
 XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma;  
 KM foetal abnormality; developmental abnormality; hematopoietic disorder;  
 KM immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 KM Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;  
 KM psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;  
 KM inflammation; neurological disorder; Alzheimer's disease; food additive;  
 KM angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;  
 KM pregnancy-related disorder; endocrine disorder; infection; wound healing;  
 KM cell culture; chemotaxis; vunerary; binding partner identification;  
 KM gene therapy.  
 KM  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX  
 XX Key Location/Qualifiers  
 FT Peptide 1..24  
 FT /label= signal\_peptide  
 FT 25..472  
 FT Protein /label= Mature\_human\_secreted\_protein  
 XX  
 XX W0200151504-A1.  
 FN  
 XX  
 XX 19-JUL-2001.  
 PD  
 XX  
 XX 12-JAN-2001; 2001WO-US00911.  
 PF  
 XX  
 XX 13-JAN-2000; 2000US-0482273.  
 PR  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX  
 XX Ruben SM, Komatsoulis GA, Duan DR, Rosen CA, Moore PA, Shi Y;  
 PI Lafleur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;  
 PI Endress GA, Muscenski M, Emner R;  
 PI  
 XX WPI; 2001-425865/45.  
 DR  
 XX  
 XX N-PSDB; AAD11647.  
 DR  
 XX

PT Isolated nucleic acid molecule encoding a human secreted protein is  
 PT used in preventing, treating or ameliorating a medical condition -  
 XX  
 PS Claim 11; Page 750-752; 864pp; English.

CC AAD1630-AAD11721 represent cDNAs corresponding to 71 human secreted  
 CC protein genes, and AAB06041-AAB06132 represent the proteins they encode.  
 CC AAB06133-AAB06205 represent human secreted protein fragments.  
 CC The secreted proteins and their genes are useful for preventing, treating  
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
 CC Pathological conditions can be diagnosed by determining the amount of the  
 CC new protein in a sample or by determining the presence of mutations in  
 CC the new genes. Specific uses are described for each of the 71 genes,  
 CC based on the tissues in which they are most highly expressed, and include  
 CC developing products for the diagnosis or treatment of proliferative  
 CC disorders, cancer, tumours, foetal and developmental abnormalities,  
 CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
 CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,  
 CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
 CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
 CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
 CC angiogenic disorders, kidney disorders, gastrointestinal disorders,  
 CC pregnancy-related disorders, endocrine disorders, and infections. The  
 CC proteins can also be used to aid wound healing and epithelial cell  
 CC proliferation, to prevent skin aging due to sunburn, to maintain organs  
 CC before transplantation, for supporting cell culture of primary tissues,  
 CC to regenerate tissues, to identify their cognate ligands or binding  
 CC partners, and in chemotaxis, and can be used as a food additive or  
 CC preservative to modify storage properties. Antibodies specific for a  
 CC protein of the invention can be used in alleviating symptoms associated  
 CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
 CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA).  
 CC The present sequence represents a human secreted protein of  
 CC the invention.

XX Sequence 472 AA;

SQ Query Match 100.0%; Score 10; DB 22; Length 472;

Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TACTIVEDAE 10  
 |||||  
 Db 231 TACTIVEDAE 240

RESULT 10  
 AAB61800  
 ID AAB61800 standard; Protein; 472 AA.

XX AC AAB61800;

XX DT 04-OCT-2002 (first entry)

XX DE Human polypeptide SEQ ID NO 154.

XX XX Human; cytosolic; antirheumatic; antiarthritic; vulnery; analgesic;  
 KM antinflammatory; antibacterial; immunosuppressive; antiparkinsonian;  
 KM neuroprotective; nociceptive; osteopathic; haemostatic; vasotrophic;  
 KM antifungal; fungicide; antidiabetic; antiallergic;  
 KM immunostimulant; antiparasitic; secreted protein; transmembrane protein;  
 KM cytokine; cell proliferation; cell differentiation; autoimmune disease;  
 KM stem cell; growth factor; nervous system disease; neuropathy;  
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease;  
 KM osteoporosis; severe combined immunodeficiency; SCID; infection;  
 KM multiple sclerosis; rheumatoid arthritis; gene therapy.

XX OS Homo sapiens.

XX PN US2002065394-A1.

XX PD 30-MAY-2002.

XX XX

PF 22-DEC-2000; 2000US-0745763.  
 XX  
 PR 18-MAR-1998; 98US-0040963.

XX (JACO/) JACOBS K.  
 PA (MCCO/) MCCOY J M.  
 PA (LAVA/) LAVALLIE E R.  
 PA (COLL/) COLLINS-RACIE L A.  
 PA (EVAN/) EVANS C.  
 PA (MERB/) MERBERG D.  
 PA (TREAC/) TREACY M.  
 PA (SPAUV/) SPAULDING V.

PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;  
 PI Metzberg D, Treacy M, Spaulding V;

XX WPI; 2002-582343/62.  
 DR N-PSDB; ABQ92016.

PT Novel secreted or transmembrane protein and polynucleotide encoding the  
 PT protein, useful for diagnosis and treatment of neurological disorders,  
 PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis  
 PT  
 PS Claim 51; Page 114-115; 284pp; English.

CC The invention relates to human secreted or transmembrane protein (I),  
 CC their fragments and is encoded by specific complementary deoxyribonucleic  
 CC acid (cDNA) inserts (II), where the protein is substantially free from  
 CC other mammalian proteins. (I) are useful for preventing, treating or  
 CC ameliorating a medical condition, especially immunological treatment or  
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,  
 CC cytokine, cell proliferation, cell differentiation, antiinflammatory,  
 CC stem cell growth factor activity and activin or inhibin-related  
 CC activities. (I) can be used to manipulate stem cells in culture to give  
 CC rise to neuroepithelial cells that can be used to augment or replace  
 CC cells damaged by illness, autoimmune disease, accidental damage or  
 CC genetic disorders. (I) induces the proliferation of neural cells and  
 CC regeneration of nerve and brain tissue and is useful for the treatment of  
 CC central and peripheral nervous system diseases and neuropathies, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic  
 CC activity, regulation of haematopoiesis and is useful for treating myeloid  
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopaenia  
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,  
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or  
 CC periodontal disease. (I) is also useful for gut protection or  
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury  
 CC in various tissues, various immune deficiencies and disorders including  
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,  
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,  
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,  
 CC such as asthma or other respiratory problems. (II) is useful to express  
 CC recombinant protein, as markers for tissues in which the corresponding  
 CC protein is preferentially expressed and in gene therapy. The present  
 CC sequence is that of a polypeptide of the invention.

XX SQ Sequence 472 AA;

Query Match 100.0%; Score 10; DB 23; Length 472;

Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TACTIVEDAE 10  
 |||||  
 Db 231 TACTIVEDAE 240

RESULT 11  
 AABG33880  
 ID AABG33880 standard; Protein; 472 AA.

XX XX

AC ABG33880;  
 XX 15-JUL-2002 (first entry)  
 XX  
 DE Human secreted protein encoded by gene 18 #1.  
 XX  
 KM Human; secreted protein; gene therapy; immunosuppressive;  
 KM antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;  
 KM vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;  
 KM vitruclide; fungicide; ophthalmological; autoimmune disease; neoplasm;  
 KM rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;  
 KM cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;  
 KM angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
 KM ocular disorder; corneal infection; wound healing; skin aging;  
 KM epithelial cell proliferation; food additive.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200226931-A2.  
 XX  
 PD 04-APR-2002.  
 XX  
 PF 24-SEP-2001; 2001WO-US29871.  
 XX  
 PR 25-SEP-2000; 2000US-234925P.  
 PR 12-JAN-2001; 2001WO-US00911.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Komatsoulis G, Duan DR, Rosen CA, Moore PA, Shi Y,  
 PI Lafleur DM, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR;  
 PI Endress GA, Mucenski M, Ebner R;  
 XX  
 DR WPI: 2002-362489/39.  
 DR N-PSDB; ABK69743.  
 XX  
 PT Novel 71 isolated secreted polypeptides and polynucleotides encoding  
 PT the polypeptides, useful for treating Huntington's disease, sepsis,  
 PT meningitis, thrombocytopenia, haemolytic anaemia, rheumatoid arthritis,  
 PT asthma  
 PT  
 PS Claim 11; Page 1231-1232; 1478pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (or its  
 CC fragment, homologue complement or allelic variant) encoding a human  
 CC secreted protein (and its fragment, domain, epitope, variant, secreted  
 CC form and species variant). Also included are a recombinant vector  
 CC comprising the nucleic acid, a recombinant host cell comprising the  
 CC vector, an antibody against the secreted protein, a recombinant host cell  
 CC that expresses the secreted protein and a method of identifying a binding  
 CC partner of the secreted protein. The nucleic acid and protein are used to  
 CC prevent, diagnose, treat or ameliorate a medical condition in e.g.  
 CC human, mice, rabbits, goats, horses, cats, dogs, chickens or sheep  
 CC for example autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
 CC ocular disorders e.g. corneal infection. Many other diseases and  
 CC disorders are listed in the specification. The polypeptides can also be  
 CC used to aid wound healing in the epithelial cell proliferation, to prevent  
 CC skin aging due to sunburn, to maintain organs before transplantation, for  
 CC supporting cell culture of primary tissues, to regenerate tissues and in  
 CC chemotaxis. The polypeptides can also be used as a food additive or  
 CC preservative to increase or decrease storage capabilities. The present  
 CC sequence represents a novel human secreted protein of the invention.  
 XX  
 SQ Sequence 472 AA;

Query Match 100.0%; Score 10; DB 23; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TACTIVEDAE 10  
 |||||  
 Db 231 TACTIVEDAE 240

RESULT 12  
 ABR47892  
 ID ABR47892 standard; Protein; 472 AA.  
 XX  
 AC ABR47892;  
 XX  
 DT 12-JUN-2003 (first entry)  
 XX  
 DE Human secreted protein, SEQ ID 783.  
 XX  
 KM Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;  
 KM vulnerary; antiinflammatory; nootropic; neuroprotective;  
 KM antiparkinsonian; gene therapy; human; cardiovascular disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200295010-A2.  
 XX  
 PD 28-NOV-2002.  
 XX  
 PF 19-MAR-2002; 2002WO-US09785.  
 XX  
 PR 21-MAR-2001; 2001US-277340P.  
 PR 19-JUL-2001; 2001US-306171P.  
 PR 13-NOV-2001; 2001US-331287P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM;  
 XX  
 DR WPI: 2003-129429/12.  
 DR  
 XX  
 PT Novel human secreted proteins, useful for detecting, preventing,  
 PT diagnosing, prognosticating, treating and/or ameliorating  
 PT cardiovascular disorders such as arrhythmia -  
 PT  
 PS Claim 13; SEQ ID 783; 1881pp; English.  
 XX  
 CC The present invention relates to novel human secreted proteins  
 CC (ABR47892-ABR48145) and their coding sequences (ACC50344-ACC50856). The  
 CC proteins and their coding sequences are useful for the preparation of a  
 CC diagnostic or pharmaceutical composition for diagnosing or treating a  
 CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,  
 CC coronary arteriosclerosis and myocardial ischaemia), neural disorders,  
 CC immune system disorders, muscular disorders, reproductive disorders,  
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,  
 CC proliferative disorders and/or cancers diseases and conditions, for  
 CC wound healing and epithelial cell proliferation, to treat inflammation or  
 CC infection, for treating thrombosis and arteriosclerosis, for treating or  
 CC preventing neural damage which occurs in neural disorders or  
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
 CC disease, to enhance bone and periodontal regeneration and aid in tissue  
 CC transplants or bone grafts, to prevent skin aging or hair loss, to  
 CC stimulate growth and differentiation of hematopoietic cells and bone  
 CC marrow cells when used in combination with other cytokines, to maintain  
 CC organs before transplantation or for supporting cell culture of primary  
 CC tissues, to increase or decrease differentiation or proliferation of  
 CC embryonic stem cells, or to modulate mammalian characteristics or  
 CC metabolism.  
 CC Note: The sequence data for this patent was published in electronic  
 CC format and is available from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 472 AA;

Query Match 100.0%; Score 10; DB 24; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TACTIVEDAE 10  
|||  
Db 231 TACTIVEDAE 240

## RESULT 13

ABR48106  
ID ABR48106 standard; Protein; 472 AA.

AC ABR48106;

DT 12-JUN-2003 (first entry)

DE Human secreted protein, SEQ ID 997.

XX Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;  
KW vulnerable; antiinflammatory; neuroprotective;  
KW antiparkinsonian; gene therapy; human; cardiovascular disorder.

XX OS Homo sapiens.

XX WO200295010-A2.

XX PD 28-NOV-2002.

XX PF 19-MAR-2002; 2002WO-US09785.

XX PR 21-MAR-2001; 2001US-277340P.

XX PR 19-JUL-2001; 2001US-306171P.

XX PR 13-NOV-2001; 2001US-331287P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX DR WPI; 2003-123429/12.

XX PT Novel human secreted proteins, useful for detecting, preventing,  
PT diagnosing, prognosticating, treating and/or ameliorating  
PT cardiovascular disorders such as arrhythmia -

XX PS Claim 13; SEQ ID 997; 1881pp; English.

XX CC The present invention relates to novel human secreted proteins  
CC (ABR47633-ABR48145) and their coding sequences (ACC50344-ACC50856). The  
CC proteins and their coding sequences are useful for the preparation of a  
CC diagnostic or pharmaceutical composition for diagnosing or treating a  
CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,  
CC coronary arteriosclerosis and myocardial ischemia), neural disorders,  
CC immune system disorders, muscular disorders, reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, renal disorders, for  
CC proliferative disorders and/or cancerous diseases and conditions, for  
CC wound healing and epithelial cell proliferation, to treat inflammation or  
CC infection, for treating thrombosis and arteriosclerosis, for treating or  
CC preventing neural damage which occurs in neuronal disorders or  
CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
CC disease, to enhance bone and periodontal regeneration and aid in tissue  
CC transplants or bone grafts, to prevent skin aging or hair loss, to  
CC stimulate growth and differentiation of haematopoietic cells and bone  
CC marrow cells when used in combination with other cytokines, to maintain  
CC organs before transplantation or for supporting cell culture of primary  
CC tissues, to increase or decrease differentiation or proliferation of  
CC embryonic stem cells, or to modulate mammalian characteristics or  
CC metabolism.  
CC Note: The sequence data for this patent was published in electronic  
CC format and is available from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 472 AA;

Query Match 100.0%; Score 10; DB 24; Length 472;  
Best Local Similarity 100.0%; Pred. No. 0.0051;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 TACTIVEDAE 10  
|||  
Db 231 TACTIVEDAE 240

## RESULT 14

ABR00152  
ID ABR00152 standard; Protein; 472 AA.

AC ABR00152;

DT 03-APR-2003 (first entry)

DE Human gene 142 encoded secreted protein HRA035, SEQ ID NO:441.

XX Human; secreted protein; digestive disorder; gastrointestinal disorder;  
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;  
KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;  
KW immune disorder; inflammation; infection; wound healing; drug screening;  
KW chromosome identification; chromosome mapping; cytostatic; gene therapy;  
KW antiinflammatory; immunosuppressive; vulnerable; chromosome 8q22.2.

XX OS Homo sapiens.

XX WO200276488-A1.

XX PD 03-OCT-2002.

XX PF 19-MAR-2002; 2002WO-US08276.

XX PR 21-MAR-2001; 2001US-277340P.

XX PR 19-JUL-2001; 2001US-306171P.

XX PR 13-NOV-2001; 2001US-331287P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX DR WPI; 2003-029900/02.

XX PS N-PSDB; ABR271331.

XX PT New human secreted proteins and nucleic acids, useful for detecting,  
PT preventing, diagnosing, prognosticating, treating and/or ameliorating  
PT e.g. gastrointestinal diseases and disorders, or cancers -  
XX PS Claim 13; Page 1029-1030; 1216pp; English.

XX CC ABR27190-ABR271478 represent cDNAs corresponding to 178 human secreted  
CC protein genes, and ABR00011-ABP00299 represent the proteins they encode.  
CC ABR271479-ABR271540 represent human secreted protein genomic fragments. The  
CC invention also encompasses antibodies specific for the secreted proteins,  
CC the use of the secreted proteins in drug screening, and recombinant  
CC vectors and host cells comprising a nucleic acid of the invention. The  
CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
CC fragments specific for the secreted proteins, and modulators or preventing  
CC activity are useful for diagnosing, treating, ameliorating or preventing  
CC digestive disorders. Such conditions include disorders of the mouth,  
CC oesophagus, stomach, small intestine, large intestine, liver, biliary  
CC tract and pancreas, and include cancers of these organs and tissues. The  
CC secreted proteins and their nucleic acids may also be used in the  
CC treatment of immune disorders, inflammation, infection,  
CC hyperproliferative disorders, and to promote wound healing. Nucleic acids  
CC of the invention may be used for chromosome identification, chromosome  
CC mapping, in gene therapy, for identifying individuals from minute  
CC biological samples, as hybridisation probes, and as molecular weight  
CC markers. The present sequence represents a human secreted protein of the  
CC invention.

XX SQ Sequence 472 AA;

Query Match 100.0%; Score 10; DB 24; Length 472;



Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TACTIVEDAE 10  
 |||||  
 DB 231 TACTIVEDAE 240

## RESULT 15

ABR00274  
 ID ABR00274 standard; Protein: 472 AA.

AC ABR00274;  
 DT 03-APR-2003 (first entry).

DE Human gene 142 encoded secreted protein HRA035, SEQ ID NO:563.

KW Human; secreted protein; digestive disorder; gastrointestinal disorder;  
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;  
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;  
 KW immune disorder; inflammation; infection; wound healing; drug screening;  
 KW chromosome identification; chromosome mapping; cytostatic; gene therapy;  
 KW antiinflammatory; immunosuppressive; vulnerary; Chromosome 8q22.2.

OS Homo sapiens.

PN WO200276488-A1.

PD 03-OCT-2002.

PF 19-MAR-2002; 2002WO-US08276.

PR 21-MAR-2001; 2001US-277340P.

PR 19-JUN-2001; 2001US-306171P.

PR 13-NOV-2001; 2001US-331287P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

DR WPI; 2003-029900/02.

DR N-PSDB; ABZ71453.

PT New human secreted proteins and nucleic acids, useful for detecting,  
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating,  
 PT e.g. gastrointestinal diseases and disorders, or cancers -

PS Claim 13; Page 1108-1109; 1216pp; English.

XX ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted  
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.  
 CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The  
 CC invention also encompasses antibodies specific for the secreted proteins,  
 CC the use of the secreted proteins in drug screening, and recombinant  
 CC vectors and host cells comprising a nucleic acid of the invention. The  
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
 CC fragments specific for the secreted proteins, and modulators of protein  
 CC activity are useful for diagnosing, treating, ameliorating or preventing  
 CC digestive disorders. Such conditions include disorders of the mouth,  
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary  
 CC tract and pancreas, and include cancers of these organs and tissues. The  
 CC secreted proteins and their nucleic acids may also be used in the  
 CC treatment of immune disorders, inflammation, infection,  
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids  
 CC of the invention may be used for chromosome identification, chromosome  
 CC mapping, in gene therapy, for identifying individuals from minute  
 CC biological samples, as hybridisation probes, and as molecular weight  
 CC markers. The present sequence represents a human secreted protein of the  
 CC invention.

XX Sequence 472 AA;  
 SQ

Query Match 100.0%; Score 10; DB 24; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 0.0051;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TACTIVEDAE 10  
 |||||  
 DB 231 TACTIVEDAE 240

Search completed: December 22, 2003, 11:56:04  
 Job time : 42 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:54:11 ; Search time 21 Seconds  
(without alignments)  
45.795 Million cell updates/sec

Title: US-09-745-763-36\_COPY\_231\_240  
Perfect score: 10  
Sequence: 1 TACTIVEDAE 10

Scoring table: OLIGO  
Gapop 60.0 , Gapept 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

PIR\_76:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60.0	157	2	T05516	hypothetical prote
2	60.0	171	2	E82459	probable acetyltra
3	60.0	187	2	E87299	hypothetical prote
4	60.0	197	2	F82451	probable acetyltra
5	60.0	268	2	JQ1473	pancreatic elastase
6	60.0	333	2	F83331	probable transcrip
7	60.0	334	2	G71684	virB11 protein (vi
8	60.0	334	2	F97748	virB11 protein (im
9	60.0	344	2	T02714	hypothetical prote
10	60.0	348	2	E83920	hypothetical prote
11	60.0	349	2	T24744	hypothetical prote
12	60.0	362	2	A47211	protein kinase BRX
13	60.0	369	2	A60041	G2+/calmodulin-de
14	60.0	371	2	S51320	mitogen-activated
15	60.0	377	2	S81823	mitogen-activated
16	60.0	393	2	S51321	mitogen-activated
17	60.0	433	2	S37790	probable serine/th
18	60.0	455	2	F82345	conserved hypotet
19	60.0	465	2	JC4793	dihydroliponamide d
20	60.0	532	2	T46649	1,3-beta-glucan sy
21	60.0	571	2	D69796	two-component sens
22	60.0	591	2	S65588	ABC-transporter St
23	60.0	610	2	C89939	DnaK protein (limp
24	60.0	611	2	D84245	DNA gyrase subunit
25	60.0	639	1	A39135	DNA topoisomerase
26	60.0	671	2	AD1666	DNA ligase homolog
27	60.0	751	2	AF1294	DNA ligase homolog
28	60.0	751	2	AB3360	nitrogen regulatio
29	60.0	772	2	A48822	protein-glutamine

30	60.0	1108	2	T17455	translation initia
31	60.0	1114	2	T14351	serine/threonine-s
32	60.0	1374	2	A71724	dna-directed RNA p
33	50.0	56	2	A90156	conserved hypotet
34	50.0	61	2	B39754	myelin basic prote
35	50.0	62	2	AD1377	hypothetical prote
36	50.0	64	1	Q1BP2L	hypothetical prote
37	50.0	68	2	T02924	acyl carrier prote
38	50.0	82	1	G64370	conserved hypotet
39	50.0	94	1	IQECAB	integration host f
40	50.0	94	2	A85617	integration host f
41	50.0	94	2	C90753	integration host f
42	50.0	94	2	AG0135	conserved hypotet
43	50.0	95	2	T49005	hypothetical prote
44	50.0	97	2	S72857	hypothetical prote
45	50.0	99	2	H82572	RNA polymerase ome

#### ALIGNMENTS

##### RESULT 1

T05516  
hypothetical protein F13M23.90 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 26-Aug-1999  
C/Accession: T05516  
R/Beyan, M.; Wedler, H.; Wedler, E.; Wandut, R.; Hohesl, J.; Mewes, H.W.; Mayer, K.F.)  
submitted to the Protein Sequence Database, February 1999  
A/Reference number: Z15419  
A/Accession: T05516  
A/Molecule type: DNA  
A/Residues: 1-157 <BR>  
A/Cross-references: EMBL:AL035523  
A/Experimental source: Cultivar Columbia; BAC clone F13M23  
C/Genetics:  
A/Map position: 4  
A/Note: F13M23.90  
C/Superfamily: Arabidopsis hypothetical protein F13M23.90

Query Match  
Best Local Similarity 60.0%; Score 6; DB 2; Length 157;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9  
Db 39 ITVEDA 44

##### RESULT 2

E82459  
probable acetyltransferase VCA0436 [imported] - Vibrio cholerae (strain N16961 serogroup  
C/Species: Vibrio cholerae  
C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C/Accession: E82459  
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gilm, M.L.; Dodson, R.J.;  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A/Reference number: AB2035; MUID:20406833; PMID:10952301  
A/Accession: E82459  
A/Status: Preliminary  
A/Molecule type: DNA  
A/Residues: 1-171 <HBI>  
A/Cross-references: GB:AB004377; GB:AB003853; NID:99657831; PIDN:AAF96342.1; GSPDB:GN001;  
A/Experimental source: serogroup O1, strain N16961, biotype El Tor  
C/Genetics:  
A/Gene: VCA0436  
A/Map position: 2  
Query Match  
Best Local Similarity 60.0%; Score 6; DB 2; Length 171;  
Matches 6; Conservative 100.0%; Pred. No. 11;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TVEDAE 10  
|||||  
Db 11 TVEDAE 16

RESULT 3  
E87299  
hypothetical protein CC0406 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Dec-2002  
C:Accession: E87299  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolos  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: E87299  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-187 <STO>  
A:Cross-references: GB:AE005673; NID:913421567; PIDN:AAK22393.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC0406  
C:Superfamily: uncharacterized conserved protein

Query Match 60.0%; Score 6; DB 2; Length 187;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9  
|||||  
Db 87 ITVEDA 92

## RESULT 4

F82451  
probable acetyltransferase VCA0505 [imported] - *Vibrio cholerae* (strain N16961 serogroup  
C:Species: *Vibrio cholerae*  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: F82451  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
charlson, D.; Ermolaeva, M.D.; Vamathavan, J.; Beas, S.; Qin, H.; Drogol, I.; Sellers, F.  
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: F82451  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-197 <HEI>  
A:Cross-references: GB:AE004381; GB:AE003853; NID:99657902; PIDN:AAF66408.1; GSPDB:GN001  
A:Experimental source: serogroup O1, strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VCA0505  
A:Map position: 2

Query Match 60.0%; Score 6; DB 2; Length 197;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TVEDAE 10  
|||||  
Db 37 TVEDAE 42

## RESULT 5

JQ1473  
pancreatic elastase (EC 3.4.21.36) IV precursor - rat  
C:Species: *Rattus norvegicus* (Norway rat)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 22-Jun-1999

C:Accession: JQ1473; S23783  
R:Kang, J.; Miegand, U.; Mueller-Hill, B.  
Gene 110, 181-187, 1992  
A:Title: Identification of cDNAs encoding two novel rat pancreatic serine proteases.  
A:Reference number: JQ1471; MUID:92165057; PMID:11537555  
A:Accession: JQ1473  
A:Molecule type: mRNA  
A:Residues: 1-268 <KAN>  
A:Cross-references: EMBL:X59014; NID:956090; PIDN:CAA41753.1; PID:956091

C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase; zymogen  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-29/Domain: activation peptide #status predicted <ACT>  
F:30-268/Product: pancreatic elastase IV #status predicted <MAT>  
F:74-121,216/Active site: His, Asp, Ser #status predicted

Query Match 60.0%; Score 6; DB 2; Length 268;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TVEDAE 10  
|||||  
Db 92 TVEDAE 97

## RESULT 6

F83331  
Probable transcription regulator PA2511 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
C:Species: *Pseudomonas aeruginosa*  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: F83331  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: F83331  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-333 <STO>  
A:Cross-references: GB:AE004678; GB:AE004091; NID:9948560; PIDN:AA05899.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA2511

Query Match 60.0%; Score 6; DB 2; Length 333;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9  
|||||  
Db 71 ITVEDA 76

## RESULT 7

G71684  
viral protein (viral) RP292 - *Rickettsia prowazekii*  
C:Species: *Rickettsia prowazekii*  
C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
C:Accession: G71684  
R:Anderson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Almark, U  
Nature 396, 133-140, 1998  
A:Title: The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria.  
A:Reference number: A71630; MUID:99039499; PMID:9823893  
A:Accession: G71684  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-334 <AND>  
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:93868717; PIDN:CAA14753.1; PID:9386085;  
A:Experimental source: strain Madrid E  
C:Genetics:

A:Gene: virB11, RP292  
C:Superfamily: tumor-inducing plasmid pTIC58 virB11 protein

Query Match 60.0%; Score 6; DB 2; Length 334;  
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9  
|||||

Db 191 ITVEDA 196

## RESULT 8

F97748 virB11 protein [imported] - Rickettsia conorii (strain Malish 7)

C:Species: Rickettsia conorii

C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 22-Oct-2001

C:Accession: F97748

R:Gata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Rd Science 293, 2093-2098, 2001

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.

A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: F97748

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-334 <STO>

A:Cross-references: GB:AE006914; PIDN:ALU02928.1; PID:G15619456; GSPDB:GN00173

C:Genetics:

A:Gene: virB11

C:Superfamily: tumor-inducing plasmid pTIC58 virB11 protein

Query Match 60.0%; Score 6; DB 2; Length 334;  
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9  
|||||

Db 191 ITVEDA 196

## RESULT 9

T02714

hypothetical protein At2g03120 [imported] - Arabidopsis thaliana

N:Alternate names: hypothetical protein T18E12.21

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 16-Feb-2001

C:Accession: T02714; E84444

R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Rot submitted to the EMBL Data Library, September 1998

A:Description: Arabidopsis thaliana chromosome II BAC T18E12 genomic sequence.

A:Reference number: Z14702

A:Accession: T02714

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-344 <ROU>

A:Cross-references: EMBL:AC005313; NID:G3548797; PID:G3548818

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.; euse, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: E84444

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-344 <STO>

A:Cross-references: GB:AE002093; NID:G3548818; PIDN:AC34490.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g03120, T18E12.21

A:Map position: 2

A:introns: 14/1; 52/3; 103/2; 132/1; 174/3; 191/3; 220/3; 244/1; 289/3; 315/3

Query Match 60.0%; Score 6; DB 2; Length 344;  
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACTIV 6  
|||||

Db 33 TACTIV 38

## RESULT 10

E83920 hypothetical protein BH2165 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C:Accession: E83920

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiran Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and s

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: E83920

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-348 <STO>

A:Cross-references: GB:AP001514; GB:BA000004; NID:G10174613; PIDN:BA05884.1; GSPDB:GN001

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH2165

Query Match 60.0%; Score 6; DB 2; Length 348;  
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9  
|||||

Db 104 ITVEDA 109

## RESULT 11

T24744 hypothetical protein T09E11.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000

C:Accession: T24744

R:McLay, K.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19930

A:Accession: T24744

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-349 <WTL>

A:Cross-references: EMBL:Z81147; PIDN:CA003538.1; GSPDB:GN00019; CESP:T09E11.11

A:Experimental source: clone T09E11

C:Genetics:

A:Gene: CESP:T09E11.11

A:Map position: 1

A:introns: 66/1; 103/1; 124/1; 191/1; 227/1; 285/3; 315/3

C:Superfamily: Caenorhabditis elegans hypothetical protein T09E5.1

Query Match 60.0%; Score 6; DB 2; Length 349;  
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9  
|||||

Db 284 ITVEDA 289

## RESULT 12

A47211 protein kinase ERK (EC 2.7.1.1-) CEK1 - yeast (Candida albicans) (fragment)

C:Species: Candida albicans

C>Date: 22-Sep-1999 #sequence\_revision 18-Nov-1994 #text\_change 10-Jul-1998

C:Accession: A47211

R:Whiteway, M.; Dignard, D.; Thomas, D.Y.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 9410-9414, 1992  
 A:Title: Dominant negative selection of heterologous genes: isolation of *Candida albicans*  
 A:Reference number: A47211; MUID:93028473; PMID:1408649  
 A:Accession: A47211  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-362 <WHI>  
 A:Note: sequence extracted from NCBI backbone (NCBI:115747)  
 C:Superfamily: kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein K  
 F:11-316/Domain: protein kinase homology <KIN>  
 F:19-27/Region: protein kinase ATP-binding motif

Query Match 60.0%; Score 6; DB 2; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVEDA 9  
 |||||  
 Db 305 ITVEDA 310

RESULT 13  
 A60041  
 C2/calcium-dependent protein kinase (EC 2.7.1.123) ERK1 - Chinese hamster (fragment  
 N:Alternate names: extracellular signal-regulated kinase 1 (ERK1); MAP kinase; microtubul  
 C:Species: *Citellus griseus* (Chinese hamster)  
 C:Date: 03-Mar-1993 #sequence\_revision 03-Mar-1993 #text\_change 17-Mar-1999  
 A:Accession: A60041  
 R:Meleche, S.; Pages, G.; Pouyssegur, J.  
 Mol. Biol. Cell 3, 63-71, 1992  
 A:Title: Functional expression and growth factor activation of an epitope-tagged p44 mit  
 A:Reference number: A60041; MUID:92199340; PMID:1372523  
 A:Accession: A60041  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-369 <MBL>  
 C:Comment: This serine/threonine kinase is activated in response to extracellular stimuli  
 C:Superfamily: kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; calmodulin binding; phosphoprotein; phosphotransferase; protein kinase  
 F:30-320/Domain: protein kinase homology <KIN>  
 F:38-46/Region: protein kinase ATP-binding motif

Query Match 60.0%; Score 6; DB 2; Length 369;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVEDA 9  
 |||||  
 Db 309 ITVEDA 314

RESULT 14  
 S51320  
 mitogen-activated protein kinase 6 (EC 2.7.1.1) - common tobacco  
 N:Alternate names: serine/threonine-specific protein kinase p43Nt66  
 C:Species: *Nicotiana tabacum* (common tobacco)  
 C:Date: 07-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 18-Jun-1999  
 A:Accession: S51320  
 R:Wilson, C.; Anglmayer, R.; Vicente, O.; Heberle-Bors, E.  
 Eur. J. Biochem. 233, 249-257, 1995  
 A:Title: Molecular cloning, functional expression in *Escherichia coli*, and characterizat  
 A:Reference number: S51320; MUID:96061956; PMID:7588752  
 A:Accession: S51320  
 A:Molecule type: mRNA  
 A:Residues: 1-371 <W12>  
 A:Cross-references: EMBL:X83879; NID:9634067; PION:CAA56760.1; PID:9634068  
 C:Genetics:  
 A:Gene: nt66  
 C:Superfamily: kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase  
 F:35-324/Domain: protein kinase homology <KIN>

F:44-52/Region: protein kinase ATP-binding motif

Query Match 60.0%; Score 6; DB 2; Length 371;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVEDA 9  
 |||||  
 Db 313 ITVEDA 318

RESULT 15  
 S48123  
 mitogen-activated protein kinase 7 (EC 2.7.1.1) - alfalfa  
 N:Alternate names: MAP kinase; MsK7 kinase  
 C:Species: *Medicago sativa* (alfalfa)  
 C:Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 18-Jun-1999  
 A:Accession: S48123; S59949  
 R:Jonak, C.; Pay, A.; Boegre, L.; Hirt, H.; Heberle-Bors, E.  
 Plant J. 3, 611-617, 1993  
 A:Title: The plant homologue of MAP kinase is expressed in a cell cycle-dependent and ori  
 A:Reference number: S48123; MUID:94035163; PMID:8220466  
 A:Accession: S48123  
 A:Molecule type: mRNA  
 A:Residues: 1-387 <JON>  
 A:Cross-references: EMBL:X66469; NID:9298018; PION:CAA47099.1; PID:9298018  
 R:Duerr, B.; Gawienowski, M.; Ropp, T.; Jacobs, T.  
 Plant Cell 5, 87-96, 1993  
 A:Title: MsERK1: a mitogen-activated protein kinase from a flowering plant.  
 A:Reference number: S59949; MUID:93177216; PMID:8439746  
 A:Accession: S59949  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-387 <DUB>  
 A:Cross-references: EMBL:L07042; NID:9289124; PION:AA841548.1; PID:9289125  
 C:Genetics:  
 A:Gene: MsK7  
 C:Superfamily: kinase-related transforming protein; protein kinase homology  
 C:Keywords: ATP; phosphotransferase; protein kinase  
 F:52-340/Domain: protein kinase homology <KIN>  
 F:61-69/Region: protein kinase ATP-binding motif

Query Match 60.0%; Score 6; DB 2; Length 387;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVEDA 9  
 |||||  
 Db 329 ITVEDA 334

Search completed: December 22, 2003, 11:57:51  
 Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:52:00 ; Search time 10 Seconds  
(without alignments)  
47.027 Million cell updates/sec

Title: US-09-745-763-36\_COPY\_231\_240  
Perfect score: 10  
Sequence: 1 TACTIVEDAE 10

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues  
Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	60.0	310	1 FMT_STRCO	Q910Y6 streptomyc
2	60.0	371	1 NTF6_TOBAC	Q40531 nicotiana t
3	60.0	387	1 NMK1_MEDSA	Q07176 medicago sa
4	60.0	393	1 NTF4_TOBAC	Q40532 nicotiana t
5	60.0	417	1 ERK1_CANAL	P28859 candida alb
6	60.0	424	1 MS11_ARATH	Q22467 arabidopsis
7	60.0	433	1 KRO1_YEAST	P36005 saccharomyc
8	60.0	532	1 G61_NEUCR	P36878 neurospora
9	60.0	610	1 DNAB_STAMU	Q98177 staphylococ
10	60.0	610	1 DNAB_STAMU	P45554 staphylococ
11	60.0	639	1 GYRB_HAILO	P21558 haloferrax s
12	60.0	772	1 ANNU_SCHAM	P52183 schistosom
13	60.0	792	1 CADB_CHICK	Q93131 gallus gall
14	60.0	1108	1 E2K3_RAT	Q92121 rattus norv
15	60.0	1114	1 E2K3_MOUSE	Q92255 mus musculu
16	60.0	1115	1 E2K3_HUMAN	Q92255 homo sapien
17	60.0	1373	1 RPOB_RICMA	Q92443 rickettsia
18	60.0	1374	1 RPOB_RICMA	Q92443 rickettsia
19	50.0	64	1 Y64_LAMB	P03773 bacterioph
20	50.0	67	1 RPO2_RALSO	Q58x90 ralsontia s
21	50.0	82	1 Y567_METUA	Q57987 methanococ
22	50.0	94	1 IHPB_ECOLI	P08756 escherichia
23	50.0	94	1 Y803_YERPE	Q84112 yersinia pe
24	50.0	99	1 RPO2_XANAC	Q84141 xanthomonas
25	50.0	99	1 RPO2_XYLYA	Q92d77 xylyella fas
26	50.0	102	1 TMOD_PSEME	Q00459 pseudomonas
27	50.0	106	1 YX83_METAC	Q84160 methanococ
28	50.0	108	1 GLPE_SALTY	Q84163 salmoneilla
29	50.0	109	1 VNS2_CVHOC	Q08853 human cor
30	50.0	121	1 GNA2_METTM	Q50787 metnanobac
31	50.0	129	1 ACPE3_HORVU	P08817 hordeum vul
32	50.0	132	1 ACPE3_HORVU	P15543 hordeum vul
33	50.0	132	1 R88_CORGL	Q84163 cornebacte

ALIGNMENTS

RESULT 1	FMT_STRCO	STANDARD;	PRT;	310 AA.
AC	Q910Y6:			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Methionyl-tRNA formyltransferase (EC 2.1.2.9).			
GN	FMT OR SC01473 OR SCL6_30C.			
OS	Streptomyces coelicolor.			
OC	Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;			
OC	Streptomycineae; Streptomycetaceae; Streptomycetes.			
OX	NCBI_TaxID=1902;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-A3(2) / M145;			
RX	MEDLINE=2196410; Pubmed=12000953;			
RA	Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,			
RA	Thomson N.R., James K.D., Harris D.B., Quail M.A., Kleser H.,			
RA	Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,			
RA	Cronin A., Fraser A., Goble A., Hidalgo J., Hornaby T., Howarth S.,			
RA	Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,			
RA	Rabinowitz E., Rajadream M.A., Rutherford K., Rutter S.,			
RA	Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,			
RA	Warren T., Wietorrek A., Woodward J., Barrrell B.G., Parkhill J.,			
RA	Hopwood D.A.;			
RT	"Complete genome sequence of the model actinomycete Streptomyces			
RT	coelicolor A3(2) ".			
RL	Nature 417:141-147(2002).			
CC	-1- FUNCTION: Modify the free amino group of the methionyl moiety of			
CC	methionyl-tRNA(fMet). The formyl group appears to play a dual role			
CC	in the initiator identity of N-formylmethionyl-tRNA by: (i)			
CC	promoting its recognition by IF2 and (ii) impairing its binding to			
CC	EF-Tu-GTP (By similarity).			
CC	-1- CATALYTIC ACTIVITY: 10-formyltetrahydrofolate + L-methionyl-			
CC	tRNA(fMet) + H(2)O = tetrahydrofolate + N-formylmethionyl-			
CC	tRNA(fMet).			
CC	-1- SIMILARITY: Belongs to the fmt family.			
CC	-----			
CC	EMBL: AL393109; CAB76895.1; -			
CC	HSSP: P23882; 1FMT.			
CC	HAMAP: MF_00182; -; 1.			
CC	InterPro: IPR005794; Fmt.			
CC	InterPro: IPR005793; Formyl_trans_C.			
CC	InterPro: IPR002376; formyl_transf.			
CC	InterPro: IPR001555; GART.			
CC	Pfam: PF02911; formyl_trans_C; 1.			
CC	Pfam: PF00551; formyl_transf; 1.			

DR TIGRFAMS: TIGR00460; fnc. 1.  
 DR PROSITE: PS00373; GART; FALSE NEG.  
 KW Transferrase; Methyltransferase; Protein biosynthesis;  
 KW Complete Proteome.  
 FT BINDING 110 113 TETRAHYDROFLATE (THF) (BY SIMILARITY).  
 SQ SEQUENCE 310 AA; 32639 MW; A2BA4098D6F21C2E CRC64;  
 Query Match 60.0%; Score 6; DB 1; Length 310;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 ITVEDA 9  
 DB 205 ITVEDA 210

RESULT 2  
 NTF6\_TOBAC STANDARD; PRT; 371 AA.  
 ID NTF6\_TOBAC  
 AC Q40531;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Mitogen-activated protein kinase homolog NTF6 (EC 2.7.1.-) (P43).  
 GN NTF6.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asterids; Lamiales; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Petit Havana SRI;  
 RX MEDLINE=96061956; PubMed=7588752;  
 RA Wilson C., Angimayer R., Vicente O., Heberle-Bors E.;  
 RT "Molecular cloning, functional expression in *Escherichia coli*, and  
 RT characterization of multiple mitogen-activated-protein kinases from  
 RT tobacco.";  
 RL Eur. J. Biochem. 233:249-257(1995).

CC -1- ENZYME REGULATION: Activated by tyrosine and threonine  
 CC phosphorylation (By similarity).  
 CC -1- PTM: VERY LOW AUTOPHOSPHORYLATION, ALTHOUGH DRAMATICALLY INCREASED  
 CC WHEN MN(2+) IS ADDED TO THE REACTION INSTEAD OF MG(2+).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE SUBFAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----  
 CC EMBL: X63879; CAA58760.1; -.  
 CC PIR: S68189; S51320.  
 CC HSSP: Q16539; 1MFC.  
 DR InterPro: IPR003527; MAP\_kin.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR002290; Ser\_thr\_kinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SMO0220; S\_TKC; 1.  
 DR PROSITE: PS01351; MAPK; 1.  
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE: PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE: PS00108; PROTEIN KINASE ST; 1.  
 KW Transferrase; Serine/threonine-protein kinase; ATP-binding;  
 KW Phosphorylation; Multigene family  
 FT DOMAIN 38 324 PROTEIN KINASE.  
 FT NP\_BIND 44 52 ATP (BY SIMILARITY).  
 FT BINDING 67 67 ATP (BY SIMILARITY).  
 FT ACT\_SITE 164 164 BY SIMILARITY.

FT MOD\_RES 196 196 PHOSPHORYLATION (ACTIVATES THE KINASE)  
 FT (BY SIMILARITY).  
 FT MOD\_RES 198 198 PHOSPHORYLATION (ACTIVATES THE KINASE)  
 FT (BY SIMILARITY).  
 SQ SEQUENCE 371 AA; 42741 MW; 4D97C41AC203C272 CRC64;  
 Query Match 60.0%; Score 6; DB 1; Length 371;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 ITVEDA 9  
 DB 313 ITVEDA 318

RESULT 3  
 MKK1\_MEDSA STANDARD; PRT; 387 AA.  
 ID MKK1\_MEDSA  
 AC Q07176;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Mitogen-activated protein kinase homolog MKK1 (EC 2.7.1.-) (MAP kinase  
 DE MSK7) (MAP kinase ERK1).  
 GN MKK1 OR MSK7 OR ERK1.  
 OS Medicago sativa (Alfalfa).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eubosids I; Fabales; Fabaceae; Papilionoideae; Trifoliales; Medicago.  
 OX NCBI\_TaxID=3879;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=94035163; PubMed=8220466;  
 RA Jonak C., Pay A., Boegre U., Hirt H., Heberle-Bors E.;  
 RT "The plant homolog of MAP kinase is expressed in a cell cycle-  
 RT dependent and organ-specific manner.";  
 RL Plant J. 3:611-617(1993).

CC -1- FUNCTION: MAY PLAY A ROLE IN THE MITOGENIC INDUCTION OF SYMBIOTIC  
 CC ROOT NODULES ON ALFALFA BY RHIZOBIUM SIGMAU MOLECULUS.  
 CC -1- ENZYME REGULATION: Activated by tyrosine and threonine  
 CC phosphorylation (By similarity).  
 CC -1- TISSUE SPECIFICITY: ROOTS AND STEMS.  
 CC -1- PTM: Autophosphorylated.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE SUBFAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----  
 CC EMBL: X66469; CAA47099.1; -.  
 CC EMBL: L07042; AAB41548.1; -.  
 CC PIR: S48123; S48123.  
 CC HSSP: P27703; 1BRK.  
 DR InterPro: IPR003527; MAP\_kin.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR002290; Ser\_thr\_kinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SMO0220; S\_TKC; 1.  
 DR PROSITE: PS01351; MAPK; 1.



```
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Mitosis; Cell cycle; Conjugation; Phosphorylation.
FT DOMAIN 55 340 PROTEIN KINASE.
FT NP_BIND 61 69 ATP (BY SIMILARITY).
FT BINDING 84 84 ATP (BY SIMILARITY).
FT ACT_SITE 181 213 BY SIMILARITY.
FT MOD_RES 213 213 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT MOD_RES 213 213 (BY SIMILARITY).
FT MOD_RES 215 215 PHOSPHORYLATION (ACTIVATES THE KINASE).
SQ SEQUENCE 387 AA; 44401 MW; 8672871AD34EAD2 CRC64;
Query Match 60.0%; Score 6; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
Db 329 ITVEDA 334

RESULT 4
NTF4_TOBAC STANDARD; PRT; 393 AA.
ID NTF4_TOBAC STANDARD; PRT; 393 AA.
AC Q40532;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Mitogen-activated protein kinase homolog NTF4 (EC 2.7.1.-) (P45).
GN NTF4.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_Taxid=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Petit Havana SRI;
RX MEDLINE=96061956; PubMed=7588752;
RA Wilson C., Angliaver R., Vicente O., Heberle-Bors E.;
RT "Molecular cloning, functional expression in Escherichia coli, and
RT characterization of multiple mitogen-activated-protein kinases from
RT tobacco."
RL Eur. J. Biochem. 233:249-257(1995).
CC -1- ENZYME REGULATION: Activated by tyrosine and threonine
CC phosphorylation (By similarity).
CC -1- PTM: VERY LOW AUTOPHOSPHORYLATION, ALTHOUGH DRAMATICALLY INCREASED
CC WHEN MN(2+) IS ADDED TO THE REACTION INSTEAD OF MG(2+).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X83880; CAAS8761.1; -.
DR PIR; S68190; S51321.
DR HSP; P27703; IERK.
DR InterPro; IPR003527; MAP_kin.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS01351; MAPK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
```

```
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Multigene family.
FT DOMAIN 60 345 PROTEIN KINASE.
FT NP_BIND 66 74 ATP (BY SIMILARITY).
FT BINDING 89 89 ATP (BY SIMILARITY).
FT ACT_SITE 186 186 BY SIMILARITY.
FT MOD_RES 218 218 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT MOD_RES 220 220 (BY SIMILARITY).
FT MOD_RES 220 220 PHOSPHORYLATION (ACTIVATES THE KINASE)
SQ SEQUENCE 393 AA; 45119 MW; 76EC1F3B1F74AB9 CRC64;
Query Match 60.0%; Score 6; DB 1; Length 393;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
Db 334 ITVEDA 339

RESULT 5
ERK1_CANAL STANDARD; PRT; 417 AA.
ID ERK1_CANAL STANDARD; PRT; 417 AA.
AC P28869; P87079; P87080; P87081; P87082; P87083; P87084; P87085;
AC P87086; P87322;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Extracellular signal-regulated kinase 1 (EC 2.7.1.-) (ERK1) (MAP
DE kinase 1) (MAPK 1).
GN ERK1 OR CEK1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_Taxid=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NO-1;
RX MEDLINE=93028473; PubMed=1409649;
RA Whiteaway M., Dignard D., Thomas D.V.;
RT "Dominant negative selection of heterologous genes: isolation of
RT Candida albicans genes that interfere with Saccharomyces cerevisiae
RT mating factor-induced cell cycle arrest."
RL Proc. Natl. Acad. Sci. U.S.A. 89:9410-9414(1992).
RN [2]
RP SEQUENCE OF 1-54 FROM N.A.
RC STRAIN=ATCC 36232, ATCC 60193, and Various R strains;
RX Metzgar D.L., Field D., Haubrich R., Wills C.;
RA Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -1- ENZYME REGULATION: Activated by tyrosine and threonine
CC phosphorylation (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M76585; AAA34343.2; -.
DR EMBL; U95784; AAB88588.1; -.
DR EMBL; U95785; AAB88589.1; -.
DR EMBL; U95786; AAB88590.1; -.
DR EMBL; U95787; AAB88591.1; -.
DR EMBL; U95788; AAB88592.1; -.
DR EMBL; U95792; AAB88596.1; -.
DR EMBL; U95793; AAB88597.1; -.
DR EMBL; U95795; AAB88599.1; -.
```

```

DR EMBL; U95796; AAB8600.1; -
DR EMBL; U95797; AAB8601.1; -
DR EMBL; U95798; AAB8602.1; -
DR InterPro; IPR003527; MAP_kin.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS01351; MAPK_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Mitosis; Cell cycle; Phosphorylation.
KW DOMAIN; 6 50 ALA/GLN-RICH.
FT DOMAIN 68 371 PROTEIN KINASE.
FT NP_BIND 74 82 ATP (BY SIMILARITY).
FT BINDING 97 97 ATP (BY SIMILARITY).
FT ACT_SITE 192 192 BY SIMILARITY.
FT MOD_RES 228 228 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT MOD_RES 230 230 (BY SIMILARITY).
FT MOD_RES 230 230 PHOSPHORYLATION (ACTIVATES THE KINASE)
FT VARIANT 33 33 A -> Q (IN STRAIN ATCC 60193; ALLELE 1D).
FT VARIANT 34 34 Q -> QAOQA (IN STRAIN R-2805; ALLELE 15D).
FT VARIANT 35 35 Q -> A (IN STRAIN R-2617; ALLELE 10D, STRAIN R-2624; ALLELE 14D AND STRAIN R-2777; ALLELE 12D).
FT VARIANT 38 38 Q -> QQQ (IN STRAIN R-2805; ALLELE 15D).
FT VARIANT 38 38 Q -> QQQQ (IN STRAIN R-2540; ALLELE 4D, STRAIN R-2624; ALLELE 14D AND STRAIN R-2777; ALLELE 12D).
FT VARIANT 38 38 Q -> QQQQQ (IN STRAIN R-2535; ALLELE 5D AND STRAIN R-2617; ALLELE 10D).
FT VARIANT 38 38 Q -> QQQQQQ (IN STRAIN R-2621; ALLELE 9D).
FT VARIANT 45 46 MISSING (IN STRAIN ATCC 36232; ALLELE 2D).
FT VARIANT 49 49 A -> T (IN STRAIN R-2617; ALLELE 10D).
FT VARIANT 49 49 A -> T (IN STRAIN R-2617; ALLELE 10D).
SQ SEQUENCE 417 AA; 48054 MW; 11D394A6F2BA872E CRC64;
Query Match 60.0%; Score 6; DB 1; Length 417;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 TVEDA 9
DB 360 ITVEDA 365

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty pl and TAC
RT clones.";
RL DNA Res. 7:31-63 (2000).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Contains 6 WD repeats.
CC -1- SIMILARITY: BELONGS TO THE RBAP46/RBAP48/MS11 FAMILY OF WD-REPEAT
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; AF016846; AAB70242.1; -
DR EMBL; AB019228; BAA96914.1; -
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR SMART; SMO0320; WD40; 6.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
DR PROSITE; PS50082; WD_REPEATS_2; 5.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat; Nuclear protein.
FT REPEAT 123 163 WD 1.
FT REPEAT 176 216 WD 2.
FT REPEAT 225 265 WD 3.
FT REPEAT 271 311 WD 4.
FT REPEAT 315 355 WD 5.
FT REPEAT 372 412 WD 6.
SQ SEQUENCE 424 AA; 48194 MW; 367DF89C861ECADF CRC64;
Query Match 60.0%; Score 6; DB 1; Length 424;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 5 TVEDAE 10
DB 355 TVEDAE 360

```

```

RESULT 6
MS11_ARATH STANDARD; PRT; 424 AA.
ID MS11_ARATH
AC O22467;
DT 15-JUN-1998 (Rel. 36, Created)
DT 15-JUN-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE WD-40 repeat protein MS11.
GN MS11 OR AT5G58220 OR MCK7.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;
OC eustosida II; Brassicales; Brassicaceae; Arabidopsis.
OC NCB1_TaxID=3702;
OX NCB1_TaxID=3702;
RP SEQUENCE FROM N.A.
RX MEDLINE=97480095; PubMed=9338962;
RA Ach R.A., Taranto P., Guissem W.;
RT "A conserved family of WD-40 proteins binds to the retinoblastoma
RL Plant Cell 9:1595-1606 (1997).
RN [2]

```

```

RESULT 7
KK01_YEAST STANDARD; PRT; 433 AA.
ID KK01_YEAST
AC P36005;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Probable serine/threonine-protein kinase YKL161C (EC 2.7.1.-).
GN YKL161C OR YKL615.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCB1_TaxID=4932;
OX NCB1_TaxID=4932;
RP SEQUENCE FROM N.A.
RX STRAIN=S288c;
RX MEDLINE=94378720; PubMed=8091859;
RA Vandenbol M., Bolle P.-A., Dion C., Portetelle D., Hilger F.;
RT "DNA sequencing of a 36.2 kb fragment located between the FAS1 and
RL Yea10:535-S40(1994).
CC -1- SIMILARITY: BELONGS TO THE SBR/THR FAMILY OF PROTEIN KINASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; Z26877; CAAB1493.1; -  
DR EMBL; Z28161; CAAB2003.1; -  
DR PIR; S37790; S37790.  
DR HSSP; P27703; IGOL.  
DR SGD; S0001644; YKL161C.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; Pkinase; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00220; S\_TKC; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;  
KW ATP-binding.  
FT DOMAIN 23 318 PROTEIN\_KINASE.  
FT NP\_BIND 29 37 ATP (BY SIMILARITY).  
FT BINDING 55 55 ATP (BY SIMILARITY).  
FT ACT\_SITE 153 153 BY SIMILARITY.  
SQ SEQUENCE 433 AA; 49632 MW; BC27B767B95822 CRC64;

Query Match 60.0%; Score 6; DB 1; Length 433;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITVEDA 9  
| | | | |  
Db 307 ITVEDA 312

## RESULT 8

GSI\_NEUCR STANDARD; PRT; 532 AA.  
AC P38678;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Glucan synthase-1 (EC 2.4.1.34) (1,3-beta-glucan synthase)  
DE (UDP-glucose-1,3-beta-D-glucan glucosyltransferase).  
GN GS-1 OR B13020.130.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;

[1]  
[1] SEQUENCE FROM N.A.  
RC STRAIN=74-OR23-1A / FGSC 987;  
RX MEDLINE=95023939; PubMed=7937796;  
RA Enderlin C.S., Seilitrennikoff C.P.;  
RT "Cloning and characterization of a Neurospora crassa gene required  
RT for (1,3) beta-glucan synthase activity and cell wall formation."  
RT Proc. Natl. Acad. Sci. U.S.A. 91:9500-9504(1994).  
RL [2]

RP SEQUENCE FROM N.A.  
RC STRAIN=74-OR23-1A / FGSC 987;  
RX PubMed=12655011;  
RA Mannhaupt G., Monttrone C., Haase D., Mewes H.-W., Aign V.,  
RA Hodelsel J.D., Fartmann B., Nyakatura G., Kempken F., Walter J.,  
RA Schulte U.;  
RT "What's in the genome of a filamentous fungus? Analysis of the  
RT Neurospora genome sequence."  
RT Nucleic Acids Res. 31:1944-1954(2003).

CC -1- FUNCTION: INVOLVED IN (1,3)BETA-GLUCAN SYNTHASE ACTIVITY AND  
CC CELL-WALL FORMATION.  
CC -1- CATALYTIC ACTIVITY: UDP-glucose + ((1,3)-beta-D-glucosyl) (N) = UDP  
CC + ((1,3)-beta-D-glucosyl) (N+1).  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; U09275; AAA50440.1; -  
DR EMBL; AL451015; CAC18203.1; -  
DR PIR; T46649; T46649.  
DR Transferase; Glycosyltransferase.  
SQ SEQUENCE 532 AA; 58939 MW; A36C00CA96F2F21B CRC64;

Query Match 60.0%; Score 6; DB 1; Length 532;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TVEDAE 10  
| | | | |  
Db 520 TVEDAE 525

## RESULT 9

DNAX\_STAM STANDARD; PRT; 610 AA.  
AC O997;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Chaperone protein dnax (Heat shock protein 70) (Heat shock 70 kDa  
DE protein) (HSP70).  
GN DNAX OR SAV1580 OR SA1409 OR MW1532.  
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),  
OS Staphylococcus aureus (strain N315), and  
OS Staphylococcus aureus (strain MW2).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=158878, 158879, 196620;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=MU50 / ATCC 700699, and N315;  
RX MEDLINE=21311952; PubMed=11418146;  
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iino T., Ito T.,  
RA Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hasegawa A.,  
RA Mizutani U.Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
RA Sekimizu K., Hiraoka H., Kuhara S., Goto S., Yabuzaki J.,  
RA Kanehisa M., Yamashta A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;  
RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
RT aureus."  
RT Lancet 357:1225-1240(2001).  
RL [2]

RP SEQUENCE FROM N.A.  
RC STRAIN=MW2;  
RX MEDLINE=22040717; PubMed=12044378;  
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,  
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,  
RA Yamamoto K., Hiramatsu K.;  
RT "Genome and virulence determinants of high virulence community-  
RT acquired MRSA."  
RT Lancet 359:1819-1827(2002).

CC -1- FUNCTION: Acts as a chaperone (By similarity).  
CC -1- INDUCTION: By stress conditions e.g. heat shock (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AP003362; BAB57742.1; -  
DR EMBL; AP003134; BAB42672.1; -  
DR EMBL; AP004827; BAB95397.1; -  
DR PIR; C89939; C89939.  
DR HSSP; P04475; 1D64.  
DR HAMAP; MF\_00332; -; 1.  
DR InterPro; IPR001023; Hsp70.  
DR Pfam; PF000012; HSP70; 1.  
DR PRINTS; PR00301; HEATSHOCK70.  
DR PRODOM; PD000089; HSP70; 1.  
DR PROSITE; PS00297; HSP70\_1; 1.  
DR PROSITE; PS00329; HSP70\_2; 1.  
DR PROSITE; PS01036; HSP70\_3; 1.  
KW Chaperone; ATP-binding; Heat shock; Phosphorylation;  
KW Complete proteome.  
FT MOD\_RES 173 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
SQ SEQUENCE 610 AA; 66361 MW; B71645C36D26AED CRC64;  
  
Query Match 60.0%; Score 6; DB 1; Length 610;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 TWEEDAE 10  
DB 595 TWEEDAE 600  
  
RESULT 10  
DNAX STAU STANDARD; PRT; 610 AA.  
ID DNAX STAU STANDARD; PRT; 610 AA.  
AC P45554;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Chaperone protein dnax (Heat shock protein 70) (Heat shock 70 kDa protein) (HSP70).  
GN DNAX.  
OS Staphylococcus aureus.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=1280;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94321356; PubMed=8045913;  
RT "Molecular cloning of two new heat shock genes related to the hsp70 genes in Staphylococcus aureus."  
RL J. Bacteriol. 176:4779-4783(1994).  
CC -1- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).  
CC -1- INDUCTION: BY stress conditions e.g. heat shock (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D30690; BAA06359.1; -  
DR HSSP; P04475; 1D64.  
DR HAMAP; MF\_00332; -; 1.  
DR InterPro; IPR001023; Hsp70.  
DR Pfam; PF000012; HSP70; 1.  
DR PRINTS; PR00301; HEATSHOCK70.  
DR PRODOM; PD000089; HSP70; 1.  
DR PROSITE; PS00297; HSP70\_1; 1.  
DR PROSITE; PS00329; HSP70\_2; 1.  
DR PROSITE; PS01036; HSP70\_3; 1.  
KW Chaperone; ATP-binding; Heat shock; Phosphorylation;  
KW MOD\_RES 173 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
SQ SEQUENCE 610 AA; 66361 MW; B71645C36D26AED CRC64;

SQ SEQUENCE 610 AA; 66347 MW; B34745C36D26AED CRC64;  
  
Query Match 60.0%; Score 6; DB 1; Length 610;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 TWEEDAE 10  
DB 595 TWEEDAE 600  
  
RESULT 11  
GRB HALSQ STANDARD; PRT; 639 AA.  
ID GRB HALSQ STANDARD; PRT; 639 AA.  
AC P21558;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-MAY-1991 (Rel. 18, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE DNA gyrase subunit B (EC 5.99.1.3).  
GN GYRB.  
OS Haloflex sp. (strain Aa 2.2).  
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Haloflex.  
OX NCBI\_TaxID=2254;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91100352; PubMed=1846146;  
RA Holmes M.L., Dyall-Smith M.L.;  
RT "Mutations in DNA gyrase result in novobiocin resistance in halophilic archaeobacteria."  
RL J. Bacteriol. 173:642-648(1991).  
CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.  
CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining of double-stranded DNA.  
CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS.  
CC -1- MISCELLANEOUS: MUTATIONS IN DNA GYRASE RESULT IN NOVOBIOCIN RESISTANCE IN HALOPHILIC ARCHAEABACTERIA.  
CC -1- SIMILARITY: Belongs to the type II topoisomerase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M38373; AAB09605.1; -  
DR HSSP; P06982; 1AD6.  
DR InterPro; IPR003594; ATPbind\_ATPase.  
DR InterPro; IPR002288; DNA\_GYRB\_C.  
DR InterPro; IPR000565; DNA\_GYRB.  
DR InterPro; IPR001241; DNA\_topoisom.  
DR InterPro; IPR006171; Toprim\_dom.  
DR Pfam; PF00204; DNA\_gyraseB\_1.  
DR Pfam; PF00986; DNA\_gyraseB\_C\_1.  
DR Pfam; PF02518; HATPase\_C\_1.  
DR Pfam; PF01751; Toprim\_1.  
DR PRINTS; PR00418; TP12FAMILY.  
DR PRODOM; PD148633; DNA\_gyraseB\_C\_1.  
DR SMART; SM00387; HATPase\_C\_1.  
DR SMART; SM00433; TOP2c\_1.  
DR TIGRFAMs; TIGR01059; gyrb\_1.  
DR PROSITE; PS00177; TOPOISOMERASE II; 1.  
KW Topoisomerase; Isomerase; ATP-binding.  
SQ SEQUENCE 639 AA; 71138 MW; 691315972C5E7506 CRC64;  
  
Query Match 60.0%; Score 6; DB 1; Length 639;  
Best Local Similarity 100.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITVEDA 9  
|||||  
Db 600 ITVEDA 605

## RESULT 12

ANNU SCHAM STANDARD; PRT; 772 AA.

AC P521B3;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Annullin (Protein-glutamyl gamma-glutamyltransferase) (EC 2.3.2.13) (Transglutaminase).  
OS Schistocerca americana (American grasshopper).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota; Neoptera; Orthopteroidea; Orthoptera; Caellifera; Acridomorpha; Acridoidea; Acrididae; Cyrtacanthacridinae; Schistocerca.  
OX NCBI\_Taxid=7009;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MBLDLINE=93050772; PubMed=1358727;  
RA Singer M.A., Hortsch M., Goodman C.S., Bentley D.;  
RT "Annullin, a protein expressed at limb segment boundaries in the grasshopper embryo, is homologous to protein cross-linking transglutaminases".  
RT Dev. Biol. 154:143-159(1992).  
CC -1- FUNCTION: PARTICIPATES IN MORPHOGENETIC ACTIVITIES OF THE CELLS, MAYBE BY STABILIZING THE MEMBRANE OR SUBCORTICAL STRUCTURES OF CELLS THAT ARE UNDER MECHANICAL STRESS. PROBABLY CATALYZES THE CROSS-LINKING OF PROTEINS AND THE CONJUGATION OF POLYAMINES TO PROTEINS.  
CC -1- CATALYTIC ACTIVITY: Protein glutamine + alkylamine = protein N(5)-alkylglutamine + NH(3).  
CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).  
CC -1- SUBCELLULAR LOCATION: INTRACELLULAR AND PERIPHERALLY ASSOCIATED WITH THE INNER LEAFLET OF THE CELL MEMBRANE, USING A FATTY ACID LINKAGE.  
CC -1- TISSUE SPECIFICITY: HAS AN ANNULAR, OR RING-LIKE EXPRESSION PATTERN IN EPITHELIAL ANNULI OF DEVELOPING LIMB SEGMENT BOUNDARY CELLS. IN EMBRYOS, IT IS SEEN IN GASTRULATING CELLS, IN CELLS SURROUNDING RAPIDLY DIVIDING NEUROBLASTS, AND IN MUSCLE PIONEER CELLS INVAGINATING TO FORM APODEMS.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSION OF THIS PROTEIN IN EMBRYOS AND LIMBS IS ASSOCIATED WITH AREAS UNDERGOING MOVEMENTS, MORPHOGENETIC REARRANGEMENTS, OR RAPID CELL DIVISION. EXPRESSION OF ANNULIN PRECEDES THE FIRST MORPHOLOGICAL SIGNS OF SEGMENTATION IN THE DEVELOPING LIMBS.  
CC -1- SIMILARITY: BELONGS TO THE TRANSGLUTAMINASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; M92291; AAA29806.1; -  
DR PIR; A48822; A48822.  
DR HSSP; P00488; 1GCU.  
DR InterPro; IPR001102; GlutransfG.  
DR InterPro; IPR002931; Tnsglutase\_like.  
DR Pfam; PF01841; Tnsglut\_core; 1.  
DR Pfam; PF00927; Tnsglut\_core; 2.  
DR Pfam; PF00868; Tnsglutamin\_N; 1.  
DR SMART; SM00460; TGC; 1.  
DR PROSITE; PS00547; TRANSGLUTAMINASES; 1.  
KW Transferase; Acyltransferase; Calcium-binding; Lipoprotein; Palmitate.  
FT LIPID 4 4 PALMITATE (POTENTIAL).  
FT LIPID 5 5 PALMITATE (POTENTIAL).  
FT LIPID

FT ACT\_SITE 400 400 BY SIMILARITY.  
FT ACT\_SITE 427 427 BY SIMILARITY.  
FT METAL 467 467 CALCIUM (BY SIMILARITY).  
FT METAL 469 469 CALCIUM (BY SIMILARITY).  
FT METAL 517 517 CALCIUM (BY SIMILARITY).  
FT METAL 522 522 CALCIUM (BY SIMILARITY).  
SQ SEQUENCE 772 AA; 85941 MW; FA5A3CE6A7CAE394 CRC64;

Query Match 60.0%; Score 6; DB 1; Length 772;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 ITVEDA 10  
|||||  
Db 127 ITVEDA 132

## RESULT 13

CADB CHICK STANDARD; PRT; 792 AA.

AC 093319;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Cadherin-11 precursor.  
GN CDH11.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
OX NCBI\_Taxid=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=White Leghorn;  
RA Wei J., Dong X.R., Topouzis S., Zimmer W.E., Broders F., Thierry J.P., Koteliansky V., Majesky M.W.;  
RT "Molecular cloning of chick cadherin 11 and its expression during smooth muscle differentiation and formation of the tunica media".  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS. THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE SORTING OF HETEROGENEOUS CELL TYPES.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (potential).  
CC -1- SIMILARITY: Contains 5 cadherin domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; AF055342; AAC33675.1; -  
DR InterPro; IPR002126; Cadherin.  
DR InterPro; IPR000233; Cadherin\_C\_term.  
DR Pfam; PF00028; cadherin; 5.  
DR Pfam; PF01049; cadherin\_C\_term; 1.  
DR PRINTS; PR00205; CADHERIN.  
DR SMART; SM00112; CA; 5.  
DR PROSITE; PS00233; CADHERIN\_1; 3.  
DR PROSITE; PS50268; CADHERIN\_2; 5.  
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat; Signal.  
FT SIGNAL 1 22 POTENTIAL.  
FT PROPEP 23 53 POTENTIAL.  
FT CHAIN 54 792 CADHERIN-11.  
FT DOMAIN 54 613 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 614 634 POTENTIAL.  
FT DOMAIN 635 792 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 54 159 CADHERIN 1.  
FT DOMAIN 160 268 CADHERIN 2.

FT DOMAIN 269 383 CADHERIN 3.  
 FT DOMAIN 384 486 CADHERIN 4.  
 FT DOMAIN 487 608 CADHERIN 5.  
 FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 792 AA; 87572 MW; 3E3488C6686731AB CRC64;

Query Match  
 Best Local Similarity 60.0%; Score 6; DB 1; Length 792;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TVEDA 9  
 Db 372 TVEDA 377

RESULT 14  
 E2K3 RAT STANDARD; PRT; 1108 AA.  
 AC 092121;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Eukaryotic translation initiation factor 2-alpha kinase 3 precursor  
 DE (EC 2.7.1.-) (PKR-like endoplasmic reticulum kinase) (Pancreatic  
 DE eIF2-alpha kinase).  
 GN EIF2AK3 OR PERK OR PERK.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RC SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RC TISSUE=Pancreatic islets;  
 RC MEDLINE=99038253; PubMed=9819435;  
 RA Shi Y., Vattum K.M., Sood R., An J., Liang J., Stramm L.E., Wek R.C.;  
 RT "Identification and characterization of pancreatic eukaryotic  
 RT initiation factor 2 alpha-subunit kinase, PERK, involved in  
 RT translational control.";  
 RT Mol. Cell. Biol. 18:7499-7509 (1998).  
 RN [2]  
 RN MUTAGENESIS OF LYS-614.  
 RP MEDLINE=99150360; PubMed=10026192;  
 RA Shi Y., An J., Liang J., Hayes S.E., Sandusky G.E., Stramm L.E.,  
 RA Yang N.N.;  
 RT "Characterization of a mutant pancreatic eIF-2alpha kinase, PERK, and  
 RT co-localization with somatostatin in islet delta cells.";  
 RT J. Biol. Chem. 274:5723-5730 (1999).  
 RN [3]  
 RP SUBUNIT.  
 RX MEDLINE=20313073; PubMed=10854322;  
 RA Bertolotti A., Zhang Y., Hendershot L.M., Harding H.P., Ron D.;  
 RT "Dynamic interaction of BiP and ER stress transducers in the  
 RT unfolded-protein response.";  
 RT Nat. Cell Biol. 2:332-332 (2000).  
 CC -1- FUNCTION: PHOSPHORYLATES THE ALPHA SUBUNIT OF EUKARYOTIC  
 CC INACTIVATION-INITIATION FACTOR 2 (EIF2), LEADING TO ITS  
 CC INACTIVATION AND THUS TO A RAPID REDUCTION OF TRANSLATIONAL  
 CC INITIATION AND REPRESSION OF GLOBAL PROTEIN SYNTHESIS. SERVES AS A  
 CC CRITICAL EFFECTOR OF UNFOLDED PROTEIN RESPONSE (UPR)-INDUCED G1  
 CC GROWTH ARREST DUE TO THE LOSS OF CYCLIN D1 (BY SIMILARITY).  
 CC -1- ENZYME REGULATION: PERTURBATION IN PROTEIN FOLDING IN THE  
 CC ENDOPLASMIC RETICULUM (ER) PROMOTES REVERSIBLE DISSOCIATION FROM  
 CC HSPAS/BIP AND OLIGOMERIZATION, RESULTING IN  
 CC TRANSAUTOPHOSPHORYLATION AND KINASE ACTIVITY INDUCTION.  
 CC -1- SUBUNIT: FORMS DIMERS WITH HSPAS/BIP IN RESTING CELLS.  
 CC OLIGOMERIZES IN ER-STRESSED CELLS.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic  
 CC reticulum.  
 CC -1- TISSUE SPECIFICITY: Ubiquitous.  
 CC -1- INDUCTION: By ER stress.  
 CC -1- DOMAIN: THE LUMENAL DOMAIN SENSES PERTURBATIONS IN PROTEIN FOLDING

CC IN THE ER, PROBABLY THROUGH REVERSIBLE INTERACTION WITH HSPAS/BIP.  
 CC -1- PTM: AUTOPHOSPHORYLATED.  
 CC -1- PTM: N-GLYCOSYLATED (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC GCN2 SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; AF096835; AAC83801.1; -.  
 DR PIR; T17455; T17455.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 2.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW Serine/threonine-protein kinase; Transferase; ATP-binding;  
 KW Translation regulation; Unfolded protein response;  
 KW Endoplasmic reticulum; Phosphorylation; Glycoprotein; Signal;  
 KW Transmembrane.  
 FT SIGNAL 1 27  
 FT CHAIN 28 1108  
 FT DOMAIN 28 506  
 FT TRANSMEM 507 527  
 FT DOMAIN 528 1108  
 FT DOMAIN 585 1069  
 FT NP\_BIND 591 599  
 FT BINDING 614 614  
 FT ACT\_SITE 929 929  
 FT DOMAIN 47 50  
 FT DOMAIN 223 228  
 FT CARBOHYD 253 253  
 FT MUTAGEN 614 614  
 SQ SEQUENCE 1108 AA; 124769 MW; B3716B1FD26ED32B CRC64;

Query Match  
 Best Local Similarity 60.0%; Score 6; DB 1; Length 1108;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TVEDA 10  
 Db 66 TVEDA 71

RESULT 15  
 E2K3 MOUSE STANDARD; PRT; 1114 AA.  
 AC 0922B5; 09CTK8; 09CWT5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Eukaryotic translation initiation factor 2-alpha kinase 3 precursor  
 DE (EC 2.7.1.-) (PKR-like endoplasmic reticulum kinase) (Pancreatic  
 DE eIF2-alpha kinase).  
 GN EIF2AK3 OR PERK OR PERK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RC SEQUENCE FROM N.A., AND MUTAGENESIS OF LYS-618.  
 RP STRAIN=NIH Swiss; TISSUE=Fibroblast;  
 RX MEDLINE=99127894; PubMed=9930704;  
 RA Harding H.P., Zhang Y., Ron D.;  
 RT "Translation and protein folding are coupled by an endoplasmic-

RT reticulum-resident kinase." ;  
RL Nature 397:271-274(1999).  
RN [2]  
RP ERRATUM.  
RA Harding H.P., Zhang Y., Ron D. ;  
RL Nature 398:90-90(1999).  
RN [3]  
RP SEQUENCE OF 1-152 AND 769-1114 FROM N.A..  
RC STRAIN=C57BL/6J ;  
RX MEDLINE=21085660 ; PubMed=11217851 ;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Komoto H., Adachi T., Fukuda S.,  
RA Aizawa K., Iawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,  
RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikolaou I., Pesele G., Quackenbush J.,  
RA Schiraldi L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Guetlinchtein S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombereers P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Wiltaker C., Wilting L.,  
RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsbaki S.,  
RA Hayashizaki Y. ;  
RT "Functional annotation of a full-length mouse cDNA collection." ;  
RL Nature 409:685-690(2001).  
RN [4]  
RP SUBUNIT.  
RX MEDLINE=20313073 ; PubMed=10854322 ;  
RA Bertolotti A., Zhang Y., Hendershot L.M., Harding H.P., Ron D. ;  
RT "Dynamic interaction of BiP and ER stress transducers in the  
RT unfolded-protein response." ;  
RL Nat. Cell Biol. 2:326-332(2000).  
RN [5]  
RP FUNCTION.  
RX MEDLINE=20524051 ; PubMed=11035797 ;  
RA Brewer J.W., Diehl J.A. ;  
RT "PERK mediates cell-cycle exit during the mammalian unfolded protein  
RT response." ;  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12625-12630(2000).  
CC -1- FUNCTION: PHOSPHORYLATES THE ALPHA SUBUNIT OF EUKARYOTIC  
CC TRANSLATION-INITIATION FACTOR 2 (EIF2), LEADING TO ITS  
CC INACTIVATION AND THUS TO A RAPID REDUCTION OF TRANSLATIONAL  
CC INITIATION AND REPRESSION OF GLOBAL PROTEIN SYNTHESIS. SERVES AS A  
CC CRITICAL EFFECTOR OF UNFOLDED PROTEIN RESPONSE (UPR)-INDUCED G1  
CC GROWTH ARREST DUE TO THE LOSS OF CYCLIN D1.  
CC -1- ENZYME REGULATION: PERTURBATION IN PROTEIN FOLDING IN THE  
CC ENDOPASMIC RETICULUM (ER) PROMOTES REVERSIBLE DISSOCIATION FROM  
CC HSP45/BIP AND OLIGOMERIZATION, RESULTING IN  
CC TRANS-AUTOPHOSPHORYLATION AND KINASE ACTIVITY INDUCTION.  
CC -1- SUBUNIT: FORMS DIMERS WITH HSP45/BIP IN RESTING CELLS.  
CC OLIGOMERIZES IN ER-STRESSED CELLS.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic  
CC reticulum.  
CC -1- TISSUE SPECIFICITY: Ubiquitous.  
CC -1- INDUCTION: By ER stress.  
CC -1- DOMAIN: THE LUMENAL DOMAIN SENSES PERTURBATIONS IN PROTEIN FOLDING  
CC IN THE ER, PROBABLY THROUGH REVERSIBLE INTERACTIONS WITH HSP45/BIP.  
CC -1- PTM: AUTO-PHOSPHORYLATED.  
CC -1- PTM: N-GLYCOSYLATED.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC GGN2 SUBFAMILY.  
CC -----  
CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION  
CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -  
CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS  
CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY  
CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL  
CC ENTITIES REQUIRES A LICENSE AGREEMENT (See <http://www.isb-sib.ch/announce/>)

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

[illegible]

```

Query Match          60.0%; Score 6; DB 1; Length 1114;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 TVEDAE 10
        |||||
Db      67 TVEDAE 72

Search completed: December 22, 2003, 11:56:29
Job time : 12 secs

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: December 22, 2003, 15:10:16 ; Search time 358 Seconds  
(without alignments)  
3559.036 Million cell updates/sec

Title: US-09-745-763-36

Perfect score: 2417  
Sequence: 1 MKFLIFAFGQVHLSICSG.....AWMAVSVVADMEEMIPRS 472

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Xgapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp  
-Q=/cgn2\_1/USPFO\_epool\_p/US09745763/runat\_22122003\_113316\_25243/app\_query.fasta\_1.647  
-DB=Geneseq\_19Jun03 -QMT=faeap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09745763 @CNC 1.1 396 @runat 22122003 113316 25243 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOUTRY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	* Query Match Length	DB ID	Description
1	2417	100.0	1851 19	AAV82779
2	2417	100.0	1851 24	AB092016
3	2417	100.0	1863 21	AA298034
4	2417	100.0	1863 22	AA011647
5	2417	100.0	1863 24	ABK69743
6	2417	100.0	1863 25	ABK69817
7	2417	100.0	1863 25	AB271453
8	2417	100.0	1884 21	AA258313
9	2417	100.0	1923 21	AA298139
10	2417	100.0	2077 25	AC050603
11	2417	100.0	2077 25	AB271331
12	2410	99.7	1767 21	AAA40493
13	2410	99.7	1778 18	AAV02296
14	2402	99.4	1860 22	AAK94491
15	2338.5	96.8	1895 22	AAH99703
16	2259.5	93.5	1784 23	AA573592
17	2116	87.5	1778 24	ABK62598
18	1414	58.5	895 21	AAA44369
19	1338	55.4	1134 21	AA298101
20	1338	55.4	1134 22	AA011714
21	1338	55.4	1134 24	ABK69810
22	1338	55.4	1134 25	AC050818
23	1338	55.4	1134 25	AB271454
24	870	36.0	609 20	AAV87487
25	771	31.9	642 22	AAK92181
26	491	20.3	642 22	AAK93447
27	491	20.3	462 18	AA197398
28	491	20.3	462 18	AA198060
29	491	20.3	462 18	AAV02139
30	479	19.8	314 20	AAV63369
31	397	16.4	441 22	AAK56831
32	351	14.5	317 20	AAK50587
33	229.5	9.5	1491 20	AAK00070
34	229.5	9.5	1491 20	AAV82521
35	229.5	9.5	1491 20	AAK00018
36	226.5	9.4	1593 22	AA513324
37	225.5	9.3	1640 21	AA245188
38	190	7.9	1346 24	ABK75046
39	188	7.8	3110 21	AA258312
40	188	7.8	3171 22	AAK93781
41	188	7.8	3771 22	AAH98667
42	183	7.6	492 22	AAK92932
43	172.5	7.1	2271 21	AAK93453
44	169	7.0	1179 17	AA11558
45	169	7.0	2256 21	AAA09458

## ALIGNMENTS

RESULT 1  
AAV82779 standard; cDNA, 1851 BP.  
AAV82779;  
25-FEB-1999 (first entry)  
Clone bu45\_2 isolated from human adult placenta cDNA library.  
Secreted protein; nutritional activity; immune stimulating; vaccine;  
suppressing activity; haematopoiesis regulating activity;  
tissue growth activity; activin; inhibin activity; chemotaxis;  
chemokine activity; haemostasis; thrombolytic activity; receptor;  
ligand; anti-inflammatory; cadherin; tumour invasion suppressor;  
tumour inhibition; gene therapy; ds.

OS Homo sapiens.  
FX MN0842739-A2.  
XX PD 01-OCT-1998.  
XX PF 20-MAR-1998; 98MO-US05653.  
XX PR 19-MAR-1998; 98US-0044466.  
XX PR 21-MAR-1997; 97US-0822167.  
XX PA (GENY ) GENETICS INST INC.  
XX PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;  
PI Racie LA, Spaulding V, Treacy M;  
XX XX WPI; 1998-609890/51.  
XX DR P-PSDB; AAW85456.  
XX PT New polynucleotides encoding secreted human proteins - derived from  
XX PT human foetal brain, adult brain, foetal kidney, placenta or adult  
XX PT pineal gland cDNA libraries.  
XX PS Claim 14; Page 69-70; 113pp; English.  
XX CC The present sequence encodes a secreted protein. The polynucleotide and  
XX CC secreted protein are predicted to have biological activities which would  
XX CC make them suitable for treating, preventing or ameliorating medical  
XX CC conditions in humans and animals, although no supporting data is given.  
XX CC Suggested activities include nutritional activity, immune stimulating  
XX CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating  
XX CC activity, tissue growth activity, activin/inhibin activity,  
XX CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,  
XX CC receptor/ligand activity, anti-inflammatory activity, cadherin/cumore  
XX CC invasion suppressor activity, and tumour inhibition activity (no data is  
XX CC given in the specification to support these activities). The  
XX CC polynucleotide is also stated to be useful for gene therapy.  
SQ Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 2,096-236 Length: 1851  
Score: 2417.00 Matches: 472  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 19 Gaps: 0  
  
US-09-745-763-36 (1-472) x AAW82779 (1-1851)  
QY 1 MetLysPheLeuIlePhePheGlyValHisLeuLeuSerLeuGlySerGly 20  
DB 99 ATGAATATCTTATCTTCGATTTTCGGTGGTTCACCTTTATCCCTGCTGGG 158  
QY 21 LysAlaIleCysLeuAsnGlyIleSerLysArgThrPheGluGluIleLysGluIle 40  
DB 159 AAAGCTATATGCAAGATGCACTCTTAAGAGCACTTTGAAGAAATTAAGAAATTA 218  
QY 41 AlaSerCysGlyAspValAlaLysAlaIleIleAsnLeuAlaValIleLysAlaGln 60  
DB 219 GCCACGTGGAGATGTTCTTAAGCAATCATCAACTGTTTATGTTAAAGCCCG 278  
QY 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
DB 279 AACAGATCTATAGAGGATGGACACTTCTGTGTGATACCTGTGACCCAGACTGAGTGC 338  
QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100  
DB 339 TCCAGAACTTGAAGAAATCCATCCAAATTATGACCAAACTGCAGCAAGATGGCTG 398  
QY 101 GluLysValHisLeuGluProValArgIleProHisIleProGluArgGlyGluGluSerAla 120  
DB 399 GAGAAAGTTCACTGGAGCCAGTGAAGATACCCCACTGGAGAGGGGAGAAAGATCAGCT 458

QY 121 ValMetLeuGluProArgIleHisLeuIleAlaIleLeuGluLysSerIleGly 140  
DB 459 GTGATCTGAGAGCCAAATTCATACATAGCCATCTGGGCTTGGCAGCAGCATGGG 518  
QY 141 ThrProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160  
DB 519 ACTCTCCAGAAAGGCAATTACGAGCAAGATTCTGGTGTGACCTCTTCGATGAATCAG 578  
QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180  
DB 579 AGAAGGCTCCAGAGCAAGAGGAATGTTGTTATTAACCACTTACATCACTAC 638  
QY 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaLysValGlyAlaLeu 200  
DB 639 TCAAGAGCGGTGCAATACCAAGCCAGGGCGGTGAGAGCTGCAGAGTGGGGCTTGG 698  
QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisIleGln 220  
DB 699 GCATCTCTCATTCGATCCGTGGCTCTCTTCATCAAGTCTCACAGCTTACAGTATTG 758  
QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
DB 759 GAATACAGAGATGGCGGCCCAAAATTCACACGCTGATTAAGGTGAGAGATGACGAA 818  
QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
DB 819 ATGATGTCAGAAATGCGCTTCATAGGATCAAAATTCATTCAGCTTAAGATGGGCA 878  
QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280  
DB 879 AAAGCTACCAAGATGATGATCTCTCAACACTGATACAGAGATCATCGGAGCAAAATAT 938  
QY 281 ProGluGlnValIleValLeuValSerGlyHisLeuAspSerTrpAspValGluGlnGlyAla 300  
DB 939 CCAAGACGTTGTACTGCTGAGTGAAGATGACAGCTGGAGAGTGGAGAGGAGTCC 998  
QY 301 MetAspAspGlyGlyGlyAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320  
DB 999 ATGATGATGCGCGGTGAGCTTTATATCAAGGAGACCTCTCACTTAATTAAGATCTT 1058  
QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrAlaGluGluGlnGly 340  
DB 1059 GGGCTGCGTCCAAAGAGACTCTCGCGGTGCTCTGCACTGGAGAAAGCAAGTGGGA 1118  
QY 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360  
DB 1119 GTTGTCCTTCCAGTATTAATCACTTACAAAGGTAATATTTCCAACTACAGCTGATG 1178  
QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380  
DB 1179 ATGAGTGTGACGAGGAACCTTCTTACCACTGGGCTGCATTCACGAGCAAGTAAG 1238  
QY 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400  
DB 1239 GCCAGGCGCATCAGAGAGAGGATATAGCTCTGCGAGCCCTCAATATACATCAGGTC 1298  
QY 401 LeuSerHisGlyGluGlyThrAspIleAsnThrProIleGlnAlaGlyValProGlyVala 420  
DB 1299 CTGAGCCATGAGAGAGGAGACACATCACTTTGAGATCAAGCTGAGTCCGAGACC 1358  
QY 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisIleSerHisGlyAspThrMet 440  
DB 1359 AGTCTACTGATGACTTATCAAGATATTTCTTCTTCATCACTCCACAGGAGACCAATG 1418  
QY 441 ThrValMetAspProLysGlnMetCAsnValAlaAlaAlaValIleProAlaValSerTyr 460  
DB 1419 ACTGTCATGATCCAAAGCAGATGAATGTCTGCTGTTGGCTGTTGTTCTTAT 1478  
QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472  
DB 1479 GTTGTTCAGACATGAGAAATGCTGCTAGTCC 1514

RESULT 2  
 AB092016  
 ID AB092016 standard; cDNA; 1851 BP.  
 XX  
 AC AB092016;  
 XX  
 DT 04-OCT-2002 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 13.  
 XX  
 XX Human; cytostatic; antirheumatic; antiarthritic; vulnary; analgesic;  
 KM antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;  
 KM neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;  
 KM antitumor; fungicide; antidiabetic; antiaspheric; antiallergic;  
 KM immunostimulant; antiparasitic; secreted protein; transmembrane protein;  
 KM cytokine; cell proliferation; cell differentiation; autoimmune disease;  
 KM stem cell; growth factor; nervous system disease; neuropathy;  
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease;  
 KM osteoporosis; severe combined immunodeficiency; SCID; infection;  
 KM multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002065394-A1.  
 XX  
 PD 30-MAY-2002.  
 XX  
 PF 22-DEC-2000; 2000US-0745763.  
 XX  
 PR 18-MAR-1998; 98US-0040963.  
 XX  
 PA (JACO/) JACOBS K.  
 PA (MCCO/) MCCOY J M.  
 PA (LAVA/) LAVALLIE E R. A.  
 PA (COLL/) COLLINS-RACIE L A.  
 PA (EVAN/) EVANS C.  
 PA (MERB/) MERBERG D.  
 PA (TREAU/) TREACY M.  
 PA (SPAUD/) SPAUDING V.  
 XX  
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Spaulding V;  
 XX  
 DR WPI: 2002-582343/62.  
 DR P-PSDB: ABP61800.  
 PT Novel secreted or transmembrane protein and polynucleotide encoding the  
 PT protein, useful for diagnosis and treatment of neurological disorders,  
 PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis  
 PT  
 XX  
 PS Claim 50; Page 113-114; 284pp; English.  
 XX  
 CC The invention relates to human secreted or transmembrane protein (I),  
 CC their fragments and is encoded by specific complementary deoxyribonucleic  
 CC acid (cDNA) inserts (II), where the protein is substantially free from  
 CC other mammalian proteins. (I) are useful for preventing, treating or  
 CC ameliorating a medical condition, especially immunological treatment or  
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,  
 CC cytokine, cell proliferation, cell differentiation, antiinflammatory,  
 CC stem cell growth factor activity and activin or inhibin-related  
 CC activities. (I) can be used to manipulate stem cells in culture to give  
 CC rise to neuroepithelial cells that can be used to augment or replace  
 CC cells damaged by illness, autoimmune disease, accidental damage or  
 CC genetic disorders. (I) induces the proliferation of neural cells and  
 CC regeneration of nerve and brain tissue and is useful for the treatment of  
 CC central and peripheral nervous system diseases and neuropathies, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic  
 CC activity, regulation of haematopoiesis and is useful for treating myeloid  
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia  
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,

CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or  
 CC periodontal disease. (I) is also useful for gut protection or  
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury  
 CC in various tissues, various immune deficiencies and disorders including  
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,  
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,  
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,  
 CC such as asthma or other respiratory problems. (II) is useful to express  
 CC recombinant protein, as markers for tissues in which the corresponding  
 CC protein is preferentially expressed and in gene therapy. The present  
 CC sequence is that of a polynucleotide of the invention.  
 XX  
 XX Sequence 1851 BP; 531 A; 413 C; 438 G; 469 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2,096-236 Length: 1851  
 Score: 2417.00 Matches: 472  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0  
 US-09-745-763-36 (1-472) x AB092016 (1-1851)  
 QY 1 MetLysPheLeuIlePheAlaPheGlyGlyValHisLeuLeuSerLeuCySerGly 20  
 DB 99 ATGAAATTCCTTATCTTCGCACTTTTCGGTGGTCTTACCTTTATCCTGTGCTGGG 158  
 QY 21 LysAlaIleCyLysAangIyIleSerLysArgThrPheGluGluIle 40  
 DB 159 AAACCTATATGCAAGATGCGATCTTAAGAGCACTTTGAAGAAATTAAGAAATA 218  
 QY 41 AlaSerCyGlyAspValAlaValAlaIleLeuLeuValAlaValGlyValAlaGln 60  
 DB 219 GCCAGCTGTGAGATGTTGTCTTAAGCAATCATCACTAGCTGTTATGTGAAGCCGAG 278  
 QY 61 AsnArgSerTyrgLysArgLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80  
 DB 279 AACGATCTCTATGAGCATTCGCACTTCCTGTTGATCTGTGAGCCAGCTAGTGGC 338  
 QY 81 SerLysAnLeuGluLysAlaIleGlnIleMetTyrgLysAnLeuGluGlnAspGlyLeu 100  
 DB 339 TCCAAAGAACTTGAAGAAAGCATCAATATATGACAAACCTGCAAGATGGGCTGG 398  
 QY 101 GlnLysValHisLeuGluProValArgIleProHisTrpGluValGlyGluGlnSerAla 120  
 DB 399 GAGAAAGTTCACTCGGAGCCAGTGAATATCCCACTGGGAGGAGGAAGATCAAGCT 458  
 QY 121 ValMetLeuGluProArgIleHisLeuValAlaIleLeuGlyLeuGlySerSerIleGly 140  
 DB 459 GTGATGCTGGAGCCAGAAATTCATTAAGATACCTCTGGGCTTTGGCAGCAGCATTTGG 518  
 QY 141 ThrProProGluGlyIleThrAlaGluValLeuValAlaThrSerPheAspGluLeuGln 160  
 DB 519 ACTCCTCAGAAAGCATTAAGCAGAAAGTTCTGTGGTGACCTCTTCGATGAACCTGAG 578  
 QY 161 ArgAlaGlaSerGluAlaValGlyLysIleValValIleValArgGlnProTyrlleAsnTy 180  
 DB 579 AGAAGGGCTCAGAGGAGGAGGAGGAGATGTTGTTTATTAACCAACTTATCATCACTAC 638  
 QY 181 SerArgThrValGlnTyrgLysThrGlnGlyValAlaGluAlaValAlaValAlaLeu 200  
 DB 639 TCAAGAGCTGCAATACCAAGCAGAGGGGCGGTGGAGCTGCCAAGTGGGGCTTTG 698  
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrgSerProHisTrpGlyIleGln 220  
 DB 699 GCATCTCTCATTCATCGTGGCTCTCTTCATCTACAGTCTCTCACACAGGATTCAG 758  
 QY 221 GlnTyrgLysArgGlyValProValIleProThrAlaCysIleThrValGluAspAlaGlu 240  
 DB 759 GAATACCAAGATGCGGCGCCAAATTCACACGCTGTATAGGTGAGGAATGCGAGAA 818  
 QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260

```

Db      819  ATGATGTCAGAGATGCTTCATGATGATCAAAATTCATTCACACTAAAGTGGGGCA 878
Qy      261  LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThiGlySerIleTyr 280
Db      879  AAGACTTCCCAAGATACGATTCCTTCAACACTGACAGATGATCAGGGGCAAAAT 938
Qy      281  ProGluGlnValIleuValSerGlyHisLeuAspSerTTPAspValGlyGlnIleVal 300
Db      939  CCAAGACAGCTTGTCTGCTGATGACATCTGACAGCTGGAGCTGGAGGAGGTGCC 998-
Qy      301  MetAspAspGlyGlyGlyAlaPheIleSerTgIuAlaLeuSerLeuIleLysAspLeu 320
Db      999  ATGATGATGGGGGAGAGCTTTATATCATGAGAGCACTCTCACTTATTAAGATCTT 1058
Qy      321  GlyLeuAaGProGlyAaGThrLeuAaGLeuValLeuTTPAlaGluGluGlnIleGly 340
Db      1059  GGGCTGCTCCAAAGAGACTCTGCGCTGCTGCTGACGACGAGAAAGAACAGTGA 1118
Qy      341  ValGlyAlaPheGlnTyrTyrGlnLeuHisIleLysValAsnIleSerAsnTyrSerLeuVal 360
Db      1119  GTTGATGCTTCCTCACTATTCATGATTCACAGGTAATATTTCCAACTACAGCTTGCTG 1178
Qy      361  MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLys 380
Db      1179  ATGAGTCTGACGAGAACCTTCTTACCCACTGGGGCTGCAATTCACCTGGGAGTGA 1238
Qy      381  AlaAlaGlnIleMerGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db      1239  GCCAGGGCCATCATGAGAGAGGTTCAGCCCTGCGACGCCCTCAATATCATCTCAGCTC 1298
Qy      401  LeuSerHisGlyGlyGlyThrAspIleAsnPheTTPileGlnAlaGlyValProGlyAla 420
Db      1299  CTGACCCATGAGAGAGGACAGACATTCATCTTTGATTCAGAGCTGGAGCTTGAGACC 1358
Qy      421  SerLeuLeuAspAspLeuTyrIleTyrPhePhePheHisSerHisGlyAspThrMet 440
Db      1359  AGTCACTTGAATGATTCATACAGATTCCTCTTCATCATCTCCACAGAGACACCATG 1418
Qy      441  ThrValMetAspProGlyGlnMetCysValAlaAlaIleValTPAlaValIleSerTyr 460
Db      1419  ACTGCATGAGATCCAAAGCAGATGAAATGTCTGCTGCTGTTGGCTGTGTTCTTAT 1478
Qy      461  ValValAlaAspMetGluGluMetLeuProAspSer 472
Db      1479  GTTGTTGACGACATGAAAGAAATCTGCTGCTGCTC 1514

RESULT 3
AAZ98034
ID      AAZ98034 standard; cDNA; 1863 BP.
AC      AAZ98034;
XX      09-MAY-2000 (first entry)
DT      09-MAY-2000 (first entry)
XX      Human secreted protein encoding nucleotide sequence SEQ ID NO:28.
DE      Human; secreted protein; diagnostic; cytostatic; immunosuppressive;
XX      antiinflammatory; nootropic; neuroprotective; antiallergic; cancer;
XX      tumour; neurodegenerative disorder; developmental abnormality; allergy;
XX      foetal deficiency; blood disorder; immune system disorder; arthritis;
XX      autoimmune disease; hepatic disease; renal disease; inflammation;
XX      Alzheimer's disease; behavioural disorder; schizophrenia; osteoporosis;
XX      infection; AIDS; spinal cord injury; transplant rejection; diabetes;
XX      asthma; sepsis; acne; psoriasis; cardiovascular disorder;
XX      reproductive disorder; gastrointestinal disorder; respiratory disorder;
XX      metabolic disorder; food additive; preservative; ss.
OS      Homo sapiens.
XX      WO200004140-A1.
XX      27-JAN-2000.
PD

```

```

XX      14-JUL-1999; 99WO-US15849.
PF      15-JUL-1998; 98US-0092921.
XX      15-JUL-1998; 98US-0092922.
PR      15-JUL-1998; 98US-0092956.
XX      15-JUL-1998; 98US-0092956.
XX      (HUMA-) HUMAN GENOME SCI INC.
XX      Ruben SM, Komatsoulis G, Duan RD, Rosen CA, Moore PA, Shi Y;
XX      Lafleur DW, Ebner R, Olsen HS, Brewer LA, Florence KA, Young PE;
XX      Mueneski M, Endress GA, Soppet DR;
XX      WPI; 2000-161126/14.
XX      P-PSDB; AAY87081.
XX      New isolated human genes, useful for diagnosis and treatment of, e.g.
XX      cancers, neurological or blood disorders
XX      Claim 1; Page 319; 494pp; English.
XX      The polynucleotide sequences given in AAZ98017 to AAZ98108 encode the
XX      human secreted proteins given in AAY87064 to AAY8723. Human secreted
XX      protein can have activities based on the tissues and cells the genes are
XX      expressed in. Examples of activities include: cytostatic;
XX      immunosuppressive; antiinflammatory; nootropic; neuroprotective; and
XX      antiallergic. The polynucleotides and their corresponding secreted
XX      polypeptides are useful for preventing, treating or ameliorating medical
XX      conditions, e.g. by protein or gene therapy. Also pathological conditions
XX      can be diagnosed by determining the amount of the new polypeptides in a
XX      sample or by determining the presence of mutations in the new
XX      polynucleotides. Human secreted protein s and their polynucleotides can
XX      be used for developing products for the diagnosis or treatment of cancer,
XX      tumours, neurodegenerative disorders, developmental abnormalities and
XX      foetal deficiencies, blood disorders, diseases of the immune system,
XX      autoimmune diseases, hepatic and renal disease, inflammation,
XX      allergies, Alzheimer's disease, behavioural disorders, schizophrenia,
XX      osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
XX      transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
XX      cardiovascular disorders, reproductive disorders, gastrointestinal
XX      disorders, respiratory disorders and metabolic disorders. The
XX      proteins or polynucleotides can also be used as food additives or
XX      preservatives. The proteins are also useful for identifying their
XX      binding partners. AAZ98008 to AAZ98016 and AAY87063 are sequence used in
XX      the exemplification of the present invention.
XX      Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;
XX      Alignment Scores:
XX      Pred. No.: 2,11e-236 Length: 1863
XX      Score: 2417.00 Matches: 472
XX      Percent Similarity: 100.00% Conservative: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      DB: 21 Gaps: 0
XX      US-09-745-763-36 (1-472) x AAZ98034 (1-1863)
Qy      1 MetLysPheLeuIlePheAlaPheGlyGlyValHisLeuSerLeuCysSerGly 20
Db      99 ATGAAATTCCTTATCTTCGCTTTCGCGGTTCACCTTTATCCCTGCTGCGG 158
Qy      21 LysAlaIleCysLysAsnGlyIleSerLysAaGThrPheGluGlnIleLysGluGlnIle 40
Db      159 AAACCTTATGCAAGATGCGATCTTGAAGAGACTTTGAAGAAATPAAAGAGAAATA 218
Qy      41 AlAserCysGlyAspValAlaLysAlaIleIleLeuAsnAlaValTyrGlyLysAlaGln 60
Db      219 GCCAGCTGTGAGATGTTGCTTAAAGCAATCATCACTAGCTGTTTATGCAAGCCAG 278
Qy      61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
Db      279 AACAGATCTTATGAGCATTTGGCACTTCTGGTGTATCTGTGTGACCAAGCTGAGTGGC 338

```

QY 81 SerLySAsnLeuGluValAlaIleGlnIleMetTyrGlnAsnLeuGlnIAspGlyLeu 100  
 DB 339 TCAGAAAGCCTTAGAAAAAGCCATCCAAATTATGATCCAAAACCTGCAGCAAGATGGGCTG 398  
 QY 101 GlnLysValHisLeuGluProValArgIleProHisTTPGluArgGlyGluGluSerAla 120  
 DB 399 GAGAAAGTTCACCTGGACCCAGTGGAAATACCCCACTGGGAGAGGGGAGAAATACAGCT 458  
 QY 121 ValMetLeuGluProArgIleHisLysTyrIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
 DB 459 GTGATGCTGGAGCCAGAAATTCATAGATAGCATCTGGGCTTGGAGAGACATGGG 518  
 QY 141 ThrProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160  
 DB 519 ACTCTCCAGAAAGGATTACAGCAAGTCTGGTGGAGACTCTTTCGATGAGACTGAG 578  
 QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValValIlyrAangInProTyrIleAsnTyr 180  
 DB 579 AGAAGGGCTCAGAAAGCAAGAGGAGATGTGTTTATACCAACTTACATCACTAC 638  
 QY 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaLysValGlyAlaLeu 200  
 DB 639 TCAAGAGCGGTGCAATACCAAGCGAGGGCGGTGGAAGCTGCCAAGGTGGGGCTTGG 698  
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisTThrGlyIleGln 220  
 DB 699 GCATCTCTCATTCGATCCGTGGCTCTTCTCCATCTACATCTCTCCACACAGGATTCAG 758  
 QY 221 GlnTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
 DB 759 GAATACCGAGATGGCGTCCCAAGATTCCAACAGCTGTATTCAGTGAAGATGCAGAA 818  
 QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
 DB 819 ATGATGTCMAAAGTGGCTTCTCATGAGTCAAAATTCATTCAGTCAAGATGGGGCA 878  
 QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIleTyr 280  
 DB 879 AAGACCTACCAAGATCTGATCTCTTCAACACTGAGAGAGATCATCTGGAGCAAAAT 938  
 QY 281 ProGluGlnValIleLeuValSerGlyHisLeuAspSerTTPAspValGlyGlnGlyAla 300  
 DB 939 CCAGAACAGGTTGATCTGTCAGTGCATCTGCAGACTGCAGATGTTGGGCAAGGTGCC 998  
 QY 301 MetAspAspGlyGlyGlyAlaPheIleSerTTPGluAlaLeuSerLeuIleLysAspLeu 320  
 DB 999 ATGATGATGGCGGTGGAGCTTATATCATGGAGACACTCTCATATTAAAGATCTT 1058  
 QY 321 GlnLeuArgProLysArgThrLeuArgLeuValLeuTTPThrAlaGluGluGlnGlyGly 340  
 DB 1059 GGGCTGCTCCAAAGAGACTCTGGGCTGGTCTCTGACTGCAGAAACAGAGTGGGA 1118  
 QY 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360  
 DB 1119 GTTGCTGCTCCAGATTATTCAGTTACCAAGGTAATATTTCCTCACTACAGCTGGTG 1178  
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLys 380  
 DB 1179 ATGAGGCTGCAGCAGCAACTTCTTACCACTGGGCTGCAGATTCACGGCAGTGAAG 1238  
 QY 381 AlaArgAlaIleMetGluGluValMetSerLeuGlnInProLeuAsnIleThrGlnVal 400  
 DB 1239 GCCAGGGCCATATGAGAGAGGATTTATGAGCTCTGCAGCCCTTCATATATCATCTAGGTC 1298  
 QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTTPIleGlnAlaGlyValProGlyAla 420  
 DB 1299 CTGAGCCATGAGAGAGGACAGACATCAACTTTGGATCCAAAGCTGGAGTCTGGAGCC 1358  
 QY 421 SerLeuLeuAspAspLeuTyrIlyTyrPhePhePheHisHisSerHisGlyAspThrMet 440  
 DB 1359 AGTCACTTATGATGACTTATACAAAGTATTTCTTCCATCACTCCACGAGAGACACCATG 1418

QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaIleValTTPAlaValValSerTyr 460  
 DB 1419 ACTGTCATGATCCAAAGCAGATGATGCTGCTGCTGTTGGCTGTTGTTCTTAT 1478  
 QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472  
 DB 1479 GTTGTCAGACATGAGAAAGAAATGCTGCTTAGGTCC 1514  
 RESULT 4  
 AAD11647 ID AAD11647 standard; cDNA; 1863 BP.  
 AC AAD11647;  
 DT 24-SEP-2001 (first entry)  
 XX  
 DE Human secreted protein-encoding gene 18 cDNA clone HRACT35, SEQ ID NO:28.  
 KW Human; secreted protein; proliferative disorder; cancer; tumour; asthma;  
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;  
 KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;  
 KW inflammation; neurological disorder; Alzheimer's disease; food additive;  
 KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;  
 KW pregnancy-related disorder; endocrine disorder; infection; wound healing;  
 KW cell culture; chemotaxis; vunerary; binding partner identification;  
 KW gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 99..1517  
 FT /\*tag= a  
 FT /product= "Human secreted protein precursor"  
 FT sig\_peptide 99..170  
 FT /\*tag= b  
 FT mat\_peptide 171..1514  
 FT /\*tag= c  
 FT /product= "Mature human secreted protein"  
 FT  
 PN MO200151504-A1.  
 PD 19-JUL-2001.  
 XX  
 PF 12-JAN-2001; 2001WO-US00911.  
 XX  
 PR 13-JAN-2000; 2000US-0482273.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, KomatsuIis GA, Duan DR, Rosen CA, Moore PA, Shi Y;  
 PI Laflaur DW, Olsen HS, Brewer LA, Florence KA, Young PE, Soppet DR;  
 PI Andrews GA, Mucenski M, Edner R;  
 XX  
 DR WPI: 2001-425865/45.  
 DR P-PSDB; AAB06058.  
 PT Isolated nucleic acid molecule encoding a human secreted protein is  
 PT used in preventing, treating or ameliorating a medical condition -  
 XX  
 PS Claim 1; Page 686; 864pp; English.  
 CC AAD11630-AAD11721 represent cDNAs corresponding to 71 human secreted  
 CC protein genes, and AAB06041-AA06132 represent the proteins they encode.  
 CC AAB06133-AA06205 represent human secreted protein fragments.  
 CC The secreted proteins and their genes are useful for preventing, treating  
 CC or ameliorating medical conditions, e.g., by protein or gene therapy.  
 CC Pathological conditions can be diagnosed by determining the amount of the  
 CC new protein in a sample or by determining the presence of mutations in  
 CC the new genes. Specific uses are described for each of the 71 genes,  
 CC based on the tissues in which they are most highly expressed, and include  
 CC developing products for the diagnosis or treatment of proliferative

disorders, cancer, tumours, foetal and developmental abnormalities,  
 haematopoietic disorders, diseases of the immune system, AIDS, autoimmune  
 diseases (e.g., rheumatoid arthritis), inflammation, allergies,  
 neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),  
 cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,  
 psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,  
 angiogenic disorders, kidney disorders, gastrointestinal disorders,  
 pregnancy-related disorders, endocrine disorders, and infections. The  
 proteins can also be used to aid wound healing and epithelial cell  
 proliferation, to prevent skin aging due to sunburn, to maintain organs  
 before transplantation, for supporting cell culture of primary tissues,  
 to regenerate tissues, to identify their cognate ligands or binding  
 partners, and in chemotaxis, and can be used as a food additive or  
 preservative to modify storage properties. Antibodies specific for a  
 protein of the invention can be used in alleviating symptoms associated  
 with the disorders mentioned above, and in diagnostic immunoassays e.g.,  
 radioimmunoassay or enzyme linked immunosorbent assay (ELISA).  
 CC The present sequence represents a human secreted protein-encoding cDNA of  
 the invention.

Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

Alignment Scores:  
 Pred. No.: 2,116-236 Length: 1863  
 Score: 2417.00 Matches: 472  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0

US-09-745-763-36 (1-472) x AAD11647 (1-1863)

1 MetLysPheLeuIlePhePheGlyValHisLeuLeuSerLeuCySerGly 20  
 99 ATGAATTCCTTATCTGCAATTTTCGGTGTTGACCTTTATCCCTGCTCTG 158  
 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluIle 40  
 159 AAAGTATATGCAAGATGCGATCTCTAAGAGCATTTTGAAGAAATATAAGAAATA 218  
 41 AlaSerCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60  
 219 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCACTGCTGTTATGTTAAACCCAG 278  
 61 AsnArgSerTyrGluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80  
 279 AACGATCTCTATGACGAGTGGCACTTCTGTTGACTGTTGACCCAGACGATGAGTGC 338  
 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100  
 339 TCCAGAACTTGAAGAAAGCCATCCAAATTATGTCACAAACCTTCAGCAGATGGGCTG 398  
 101 GlnLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluGlnSerAla 120  
 399 GAGAAAGTTCCACTGAGCCAGTGAATATACCCCACTGGGAGAGGGAGAAATCAGCT 458  
 121 ValMetLeuGluProArgIleHisLysValIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
 455 GTGATGCTGAGAGCCAGAAATTCATTAAGATAGCCATCTGGGCTTGGAGCAGCACTTGG 518  
 141 ThrProGluGlyIleThrAlaGluValLeuValAlaThrSerPheAspGluLeuGln 160  
 519 ACTCTCCAGAAAGCATTAACGACAAAGTCTGGTGGTGAACCTTTTCATGATGAATGGAG 578  
 161 ArgArgAlaSerGluAlaArgGlyLysIleValAlaTyrAsnGlnProTyrIleAsnTyr 180  
 579 AGAAGGGCTCAGAGAGCAAGAGGAGATTTGTTTATACCAACCTTACATCAATCAC 638  
 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValAlaGluAlaLysValGlyAlaLeu 200  
 639 TCAGAGAGCGTGCATTAACGAAAGGAGGGGCGGTGAGAGCTGCCAAGGTGGGGCTTTG 698  
 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220

|||||  
 699 GCATCTCTCATTCGATCGGTGGCTCTCTTCATCTCAAGTCTCCACAGGATTCAG 758  
 221 GlnTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
 759 GAATACCAAGATGGCGTGGCCCAAGATTCACACCTGTATATACGGTGGAAAGATGCAGAA 818  
 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
 819 ATGATGTCAGAAATGGCTTCTCATGGGATCAAAATTTGATTCATTCAGTAAAGATGGGCA 878  
 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280  
 879 AAGACTTACCAAGATGATGATTTCTTCAACACTGATACAGAGATCATGGGAGCAAAATAT 938  
 281 ProGluGlnValValLeuValAspSerGlyHisLeuAspSerTyrPaspValGlyGlnGlyAla 300  
 939 CCAGAACAGGTTGATCGTGTGATGACATCTGACAGCTGGAGTGGTGGCAGGAGTGC 998  
 301 MetAspAspGlyGlyValAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320  
 999 ATGATGATGGCGGTGGAGCCTTTATATCATGGAGACACTCTCATTTATTAAGATCTT 1058  
 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrPThrAlaGluGlnGlyGly 340  
 1059 GGGCTGGGTCCAAAGAGAGACTCTGGCTGGTCTGTGACTGACAGAAACAGAGTGA 1118  
 341 ValGlnLysPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360  
 1119 GTTGGTGGCTTCCAGTATTCACATTCACACAGGTAATATTTCCAACTGACAGTGGTGG 1178  
 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380  
 1179 ATGAGTCTGACGAGAGACCTTTCTTACCACTGGCTGCAATTCATCGGAGTGAAG 1238  
 381 AlaArgAlaIleMetGluGlnValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400  
 1239 GCCAGGGCCATCATGAGAGAGATTATGAGCCTGCTGACGCCCTCAATATCATCACTGAGTGC 1298  
 401 LeuSerHisGlyGluGlyThrAspIleAsnPheThrIleGlnAlaGlyValProGlyAla 420  
 1299 CTGAGCCATGAGAGAGAGACATCACTTGTGATTCACAACTGGAGTGGCTGGAGCC 1358  
 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440  
 1359 AGTCTACTGATGACTTATACAGTATTTCTTCTTCATCACTCCACGAGAGACCATG 1418  
 441 ThrValMetAspProLysGlnMetAsnValAlaAlaIleValTPAlaValAlaSerTyr 460  
 1419 ACTGTCATGATCCAAAGCAGATGATGTTGCTGCTGCTGTTGGCTGTTGTTCTTAT 1478  
 461 ValValAlaAspMetGluGluMetLeuProArgSer 472  
 1479 GTTGTTCAGACATGAGAAAGATGCTGCTTAAAGTGC 1514  
 1514  
 RESULT 5  
 ABK69743 standard; cDNA; 1863 BP.  
 ID ABK69743;  
 AC ABK69743;  
 DT 15-JUL-2002 (first entry)  
 XX  
 XX Human secreted protein gene 18 #1.  
 XX  
 XX Human; ss; gene; secreted protein; gene therapy; immunosuppressive;  
 KW anti-atheritic; antihemematic; antiproliferative; cytostatic; cardiac;  
 KW vasotrophic; cerebroprotective; nootropic; neuroprotective; antibacterial;  
 KW vitruicide; fungicide; ophthalmological; autoimmune disease; neoplasm;  
 KW rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;  
 KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;  
 KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;  
 KW ocular disorder; corneal infection; wound healing; skin aging;

KW epithelial cell proliferation; food additive.  
 XX Homo sapiens.  
 XX MO200226931-A2.  
 XX 04-APR-2002.  
 XX 24-SEP-2001; 2001WO-US29871.  
 XX 25-SEP-2000; 2000US-234925P.  
 XX 12-JAN-2001; 2001WO-US00911.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Ruben SM, Komatsuji S, Duan DR, Rosen CA, Moore PA, Shi Y,  
 PI Lafleur DM, Olsen H, Brewer LA, Florence KA, Young PE, Soppet DR,  
 PI Entress GA, Mucenski W, Ebner R;  
 XX WPI; 2002-362489/39.  
 DR P-PSDB; ABG33880.  
 XX  
 PT Novel 71 isolated secreted polypeptides and polynucleotides encoding  
 PT the polypeptides, useful for treating Huntington's disease, sepsis,  
 PT meningitis, thrombocytopaenia, haemolytic anaemia, rheumatoid arthritis,  
 PT asthma  
 XX  
 PS Claim 1; Page 1170; 1478bp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (or its  
 CC fragment, homologue complement or allelic variant) encoding a human  
 CC secreted protein (and its fragment, domain, epitope, variant, secreted  
 CC form and species variant). Also included are a recombinant vector  
 CC comprising the nucleic acid, a recombinant host cell comprising the  
 CC vector, an antibody against the secreted protein, a recombinant host cell  
 CC that expresses the secreted protein and a method of identifying a binding  
 CC partner of the secreted protein. The nucleic acid and protein are used to  
 CC prevent, diagnose, treat or ameliorate a medical condition in e.g.  
 CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep  
 CC for example autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
 CC ocular disorders e.g. corneal infection. Many other diseases and  
 CC disorders are listed in the specification. The polypeptides can also be  
 CC used to aid wound healing in epithelial cell proliferation, to prevent  
 CC skin aging due to sunburn, to maintain organs before transplantation, for  
 CC supporting cell culture of primary tissues, to regenerate tissues and in  
 CC chemotaxis. The polypeptides can also be used as a food additive or  
 CC preservative to increase or decrease storage capabilities. The present  
 CC sequence encodes a novel human secreted protein of the invention.  
 XX  
 SQ Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 2,11e-236 Length: 1863  
 Score: 2417.00 Matches: 472  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0  
 US-09-745-763-36 (1-472) x ABK69743 (1-1863)  
 QY 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCySerGly 20  
 DB 99 ATGAATTCCTTATCTTGCAATTTTCGGTGGTGTACCTTTATCCCTGGCTGGG 158  
 QY 21 LysAlaIleCyLysAsnGlyYIleSerLysArgThrPheGluGluIleLysGluGluIle 40  
 DB 159 AAAGCTATATGCGAAGATGGCATCTCTAAGAGCACTTTGAAGAAATAAAGAAATA 218

QY 41 AlaSerCySerLysPheValAlaIleValAlaIleIleLeuLeuAlaValIleGlyLysAlaGln 60  
 DB 219 GCCAGCTGTGGAGATGTGTGCTTAAAGCATCATCACTTACTGTTATGTTAAAGCCAG 278  
 QY 61 AsnArgSerTyrgluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
 DB 279 AACAGATCTATAGCAGATGGACATTCCTGCTGTGATGATCTGTGGACCAAGCTAGTGGC 338  
 QY 81 SerLysAsnLeuGluLysValAlaIleGlnIleMetCysTrpGlnAsnLeuGlnIleAspGlyLeu 100  
 DB 339 TCCAAAGACCTAGAAAAAGCCATCCAAATTAATGACCAAAACCTGCAGCAAGATGGGCTG 398  
 QY 101 GluLysValHisLeuGluProValArgIleProHisTrpGluArgGlyGluGluSerAla 120  
 DB 399 GAGAAATTCACCTGGAGCCAGTGAATATCCCACTGGAGAGGGAGAAATTCAGCT 458  
 QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
 DB 459 GTGATGCTGAGGCCAAGAAATTCATTAAGATAGCCATCCTGGGCTTGGCAGCAGCATGGG 518  
 QY 141 ThrProGluGluIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160  
 DB 519 ACTCTCCAGAAAGCATTTACAGCAGAAATTCGTGTGTGTGACCTTTTCATGAACTGCAG 578  
 QY 161 ArgArgLaseGluAlaArgGlyLysIleValValTyraGlnProTyrlleAsnTyr 180  
 DB 579 AGAAGGCTCAGAGCAAGCAAGAGGAAATTCCTGTTTAAACCACTTACATCACTAC 638  
 QY 181 SerArgThrValGlnTyrArgThrGlnIleValAlaGluAlaLysValGlyAlaLeu 200  
 DB 639 TCAAGAGCGGTGCATATCCCAACGAGGGGGCGGTGAGACCTGCAGAGTTGGGGCTTTG 698  
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisTrpGlyIleGln 220  
 DB 699 GCATCTCTCATTCATCCATCCGCGCTCTCTTCATCTACAGTCCACAGGATTCAG 758  
 QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
 DB 759 GAATACACGATGGCGTCCCAAGATTCCAAGCCTGTATTACGTTGGAGAGATGCAGAA 818  
 QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
 DB 819 ATGATGTCAGAAATAGGCTTCATCAGGATCAAAATTCATTCAGTTAAGATGGGGCA 878  
 QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280  
 DB 879 AAGACCTACCCAGATCTGATTCCTTCACACTGTACAGAGATCACTGGAGAGCAAAATAT 938  
 QY 281 ProGluGlnValValIleValSerGlyHisLeuAspSerTrpAspValGlyGlnIleVala 300  
 DB 939 CCAAGACAGGTTGACTGTGTCAGTGCACATCTGCACAGCTGGAGTGTGGCAGGGTGC 998  
 QY 301 MetAspAspGlyGlyValAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320  
 DB 999 ATGATGATGGCGGTGAGCCTTATATCATGGGAACACCTCATTTAAAGATCTT 1058  
 QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrAlaGluGluGlnIleGly 340  
 DB 1059 GGGCTGTGGTCCAAAGAGACCTGCGGCTGTGTCTGAGCATGCAGAAAGAAAGGTGGA 1118  
 QY 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360  
 DB 1119 GTTGATGCTTCCAGTATATCATGATTCACAGGATTAATTTTCCAACTACAGTGGTGG 1178  
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380  
 DB 1179 ATGAGATCTGACGAGAGAACCTTCTTACCACTGGGCTGCAGATTCATCGGAGAGAAAG 1238  
 QY 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400  
 DB 1239 GCCAGGGGCATCATGAGGAGGAGTATAGCCTGTGCAGGCCCTCCAAATATCACCTCAGTGC 1298  
 QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420



Db 1299 CTGAGCCATGAGAGGAGACATCACTTTTGATCCAACTGAGTGCCTGAGACC 1358  
 Qy 421 SerLeuLeuAaPpLeuLeuTyrTyrPhePhePheHisSerHisGlyAspPheMet 440  
 Db 1359 AGTCTACTTGATGAGCTTATACAGATATTTCTTCTTCATCACTCCACGAGAACCCATG 1418  
 Qy 441 ThrValMetAaPpProGlyGlnMetAsnValAlaAlaValTTPAlaValValSerTyr 460  
 Db 1419 ACTGTCATGATGCCAAGACAGATGATGTGCTGCTGCTGTTGGGCTGTCTTTCTTAT 1478  
 Qy 461 ValValAlaAaPpMetGlnGlnMetLeuProArgSer 472  
 Db 1479 GTTGTTCAGACATGAGAAATGCTGCTAGTCC 1514  
 RESULT 6  
 ACC50817  
 ID ACC50817 standard; cDNA; 1863 BP.  
 AC ACC50817;  
 XX  
 DT 12-JUN-2003 (first entry)  
 XX  
 DE Human secreted protein coding sequence, SEQ ID 484.  
 XX  
 KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cyostatic;  
 KW vulnery; antiinflammatory; nootropic; neuroprotective;  
 KW antiparkinsonian; gene therapy; human; cardiovascular disorder;  
 KW gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PM WO200295010-A2.  
 PD 28-NOV-2002.  
 XX  
 PF 19-MAR-2002; 2002WO-US09785.  
 XX  
 PR 21-MAR-2001; 2001US-277340P.  
 PR 19-JUL-2001; 2001US-306171P.  
 PR 13-NOV-2001; 2001US-331287P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM;  
 DR WPI; 2003-129429/12.  
 PT Novel human secreted proteins, useful for detecting, preventing,  
 PT diagnosing, prognosticating, treating and/or ameliorating  
 PT cardiovascular disorders such as arrhythmia -  
 XX  
 PS Claim 21; SEQ ID 484, 1861bp; English.  
 XX  
 CC The present invention relates to novel human secreted proteins  
 CC (AB474733-AB48145) and their coding sequences (ACC50344-ACC50856). The  
 CC proteins and their coding sequences are useful for the preparation of a  
 CC diagnostic or pharmaceutical composition for diagnosing or treating a  
 CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,  
 CC coronary arteriosclerosis and myocardial ischemia), neural disorders,  
 CC immune system disorders, muscular disorders, reproductive disorders,  
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,  
 CC proliferative disorders and/or cancerous diseases and conditions, for  
 CC wound healing and epithelial cell proliferation, to treat inflammation or  
 CC infection, for treating thrombosis and arteriosclerosis, for treating or  
 CC preventing neural damage which occurs in neuronal disorders or  
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
 CC disease, to enhance bone and periodontal regeneration and aid in tissue  
 CC transplants or bone grafts, to prevent skin aging or hair loss, to  
 CC stimulate growth and differentiation of haematopoietic cells and bone  
 CC marrow cells when used in combination with other cytokines, to maintain  
 CC organs before transplantation or for supporting cell culture of primary  
 CC tissues, to increase or decrease differentiation or proliferation of

CC embryonic stem cells, or to modulate mammalian characteristics or  
 CC metabolism.  
 CC Note: The sequence data for this patent was published in electronic  
 CC format and is available from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequence.  
 XX  
 SQ Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2,11e-236 Length: 1863  
 Score: 2417.00 Matches: 472  
 Percent Similarity: 100.00% Conserved: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0  
 DB: 25  
 US-09-745-763-36 (1-472) x ACC50817 (1-1863)  
 Qy 1 MetLysPheLeuLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGly 20  
 Db 99 ATGAATTCCTTATCTTGGCATTTTTCGGTGGTTCACCTTTATCCCTGCTGCGG 158  
 Qy 21 LysAlaIleCySerLysAsnGlyIleSerLysArgThrPheGluGluIleValGluGlu 40  
 Db 159 AAAGCTATATGCAAGAAATGGCATCTCTAAGAGGACTTTTGAGAAATAAAGAAATA 218  
 Qy 41 AlaSerCySerGlyAspValAlaValAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60  
 Db 219 GCCAGCTGTGAGATGTGTGCTTAAGACATCACTCAAGCTGTTTATGTGTAAGCCCG 278  
 Qy 61 AsnArgSerTyrGluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80  
 Db 279 AACAGATCTTATGAGCGCATTTGGCACTTCTGTTATATCTGTGAGCCAGCATGAGTGC 338  
 Qy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100  
 Db 339 TCCAAGAACTTAGAAAAAGCCATCAAAATTAATGACCAAACTCGACGCAAGATGGCTG 398  
 Qy 101 GlnLysValHisLeuGluProValArgIleProHisTTPGluArgGlyGluGluSerAla 120  
 Db 399 GAGAAAGTTCACCTGGAGGCCGTAGAAATACCCACTGGAGAGGAGAAATACACT 458  
 Qy 121 ValMetLeuGluProArgGlyIleHisLysAlaIleLeuGlyLeuGlySerSerIleGly 140  
 Db 459 GTGATGCTGAGGCCAAGATTCATTAAGATACCAATCCCTGGCTTGGCAGCAGCATGGG 518  
 Qy 141 ThrProGluGlyIleThrAlaGluValLeuValThrSerPheAspGluLeuGln 160  
 Db 519 ACTCCTCCAGAGGACATTACAGAAAGTTCTGTGTGGTGAACCTTTCGATGAACCTGCG 578  
 Qy 161 ArgArgLysSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180  
 Db 579 AGAAGGGCTCAGAAAGCAAGAGGAGATGTGTGTTATACCAACTTATCATCACTAC 638  
 Qy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaIleLysValGlyAlaLeu 200  
 Db 639 TCAAGAGCGTGAATCCAGAACCCAGGCGGCGGAGACTGCAAGTTGGGCGCTTGG 698  
 Qy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisTyrIleGln 220  
 Db 699 GCATCTCTCATTCATTCATCGTGGCTCTCTTCATCTCACTCACTCTCCACAGGATTCAG 758  
 Qy 221 GlnLysArgAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
 Db 759 GAATPACAGATGGCGTCCCAAGATTCACACCTGATATTACGTGGAGATGACAGAA 818  
 Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
 Db 819 ATGATGTCAAGAAATGGCTTCATGAGATCAAAATGTGCATTCACTCAAGATGGGCGCA 878  
 Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280  
 Db 879 AAGACCTACCCAGATGATGATCTCTTCAACACTGTAGCAGATCATCTGGAGGAGAAATAT 938



QY 281 ProGluGlnValIleuValSerGlyHISLeuAspSerTrpAspValGlyGlnGlyAla 300  
 DB 999 CCAAGACAGGTTGTAAGTCTGTCAGTGACATCTGCACACTGGAGACTTGGCAGAGGTGCC 998  
 QY 301 MetAspAspGlyGlyGlyAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320  
 DB 999 ATGGATGATGGCGGGAGAGCTTTATATCATGGAGACCTCTCACTTATTAAGATCTT 1058  
 QY 321 GlyLeuAspProLysArgThrLeuArgLeuValLeuTrpThrAlaGluGlnGlyGly 340  
 DB 1059 GGGGCTGCTCCAAAGAGACTCTGGGGCTGGCTCTGCAGCTGCAGAAAGCAAGGTGGA 1118  
 QY 341 ValGlyAlaPheGlnTrpGlnLeuHisLysValAsnIleSerAsnTrpSerLeuVal 360  
 DB 1119 GTTGCTGCTTCCAGTATATTCAGTTACACANAGTAATATTTCCAACTACAGCTGGGTG 1178  
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380  
 DB 1179 ATGGAGTCTGACGAGAACCTTCTTACCCTGGGCTGCATTCACTGGCAGTGAAG 1238  
 QY 381 AlaArgAlaIleMetGluGlnValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400  
 DB 1239 GCCAGGGCCATCATGAGAGAGGTTATVAGCCCTGCAGAGCCCTCAATATCATCTCAGGTC 1298  
 QY 401 LeuSerHisGlyGlyGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420  
 DB 1299 CTGACCCATGAGAGAGAGAGACACATCACTTTGGATCCAGCTGGAGTGGCTTGAGACC 1358  
 QY 421 SerLeuLeuAspAspLeuTrpLysTrpPhePhePheHisSerHisGlyAspThrMet 440  
 DB 1359 AGTCACTTGAATGACTTATACAGATATTTCTTCCATCACTCCACAGAGACACCATG 1418  
 QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaAlaValTPAlaValAspTrp 460  
 DB 1419 ACTGCATGATGCCAAAGCAGATGATGTGCTGCTGTTGGGCTGTTGTTCTTAT 1478  
 QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472  
 DB 1479 GTTGTTCAGACATGAGAAAGAAATGCTGCTAGGTCC 1514

## RESULT 7

AB271453  
 ID AB271453 standard; cDNA, 1863 BP.

AC AB271453;

DT 04-APR-2003 (first entry)

DE Secreted protein-encoding gene 142 cDNA clone HRACU35, SEQ ID NO:274.

KW Human; secreted protein; digestive disorder; gastrointestinal disorder;  
 mouth; oesophagus; stomach; small intestine; large intestine; liver;  
 biliary tract; pancreas; cancer; tumor; hyperproliferative disorder;  
 immune disorder; inflammation; infection; wound healing; drug screening;  
 chromosome identification; chromosome mapping; cytostatic; gene therapy;  
 antiinflammatory; immunosuppressive; veterinary; chromosome 8q22.2;

KW gene; ss.

OS Homo sapiens.

PN WO200276488-A1.

PD 03-OCT-2002.

PF 19-MAR-2002; 2002WO-US08276.

PR 21-MAR-2001; 2001US-277340P.

PR 19-JUL-2001; 2001US-306171P.

PR 13-NOV-2001; 2001US-331287P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;  
 XX WPI: 2003-029900/02.  
 DR P-PSDB; ABR00274.  
 XX  
 PT New human secreted proteins and nucleic acids, useful for detecting,  
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating  
 PT e.g. gastrointestinal diseases and disorders, or cancers -  
 XX  
 PS Claim 21; Page 918, 1216pp; English.

SQ Sequence 1863 BP; 533 A; 417 C; 443 G; 470 T; 0 other;

## Alignment Scores:

Pred. No.: 2,11e-236 Length: 1863  
 Score: 2417.00 Matches: 472  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0

US-09-745-763-36 (1-472) x AB271453 (1-1863)

QY 1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCysSerGly 20  
 DB 99 ATGAAATTCCTTATCTTCGCAATTTTGGTGGTTCACCTTTATCCCTGCTGGG 158  
 QY 21 LysAlaIleCysLysAsnGlyLysSerLysArgThrPheGluGluIleLysGluGlnIle 40  
 DB 159 AAGCTATATGCAAGAAATGGATCTTAAAGGACCTTTGAAGAAATPAAAGAAATA 218  
 QY 41 AlaSerCysGlyAspValAlaLysAlaIleLeuAsnLeuAlaValTrpGlyLysAlaGln 60  
 DB 219 GCCAGCTGTGAGATGTGCTAAAGCATATCACTAGCTGTTTGTGTAAGCCAG 278  
 QY 61 AsnArgSerTrpGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
 DB 279 AACGATCTTATGAGCATGGCACTTCTGTTATATCTGTGAGACCAGACTAGTGGC 338  
 QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTrpGlnAsnLeuGlnAspGlyLeu 100  
 DB 339 TCAGAACTTAAAGAAAGCCATCAATATGTAACCAAACTGCAGCAAGAGTGGCTG 398  
 QY 101 GluLysValHisLeuGluProValArgIleProHisTrpGluArgGlyGluGluSerAla 120  
 DB 399 GAGAAAGTTACCTGAGAGCCAGTGAATATCCCACTGGAGAGAGGAGAAATGAGCT 458  
 QY 121 ValMetLeuGluProArgGlnHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
 DB 459 GTGATGCTGAGCCCAAAATTCATAGATAGCCATCTGCGGTCTTGGCAGCAGCATTTGG 518  
 QY 141 ThrProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160



Db 209 AAACCTATGCAAGATGCGATCTCTAAGAGGACCTTTGAAGAAATTAAGAAATA 268  
 Qy 41 AlaserCysglYaspValAlaIysAlaIleIleAsnLeuAlaValTYrGlyValAlaGln 60  
 Db 269 GCCAGCTGTGGAGATGTTGGCTTAAGCAATCAATCAAGCTGTTTATGTTAAGCCAG 328  
 Qy 61 AsnArgSerTYrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
 Db 329 AACAGATCTATGAGCGATTTGGCACTTCTGGTTGTAATGTGGACCCAGATGAGTGGC 388  
 Qy 81 SerIysAsnLeuGlyValAlaIleGlnIleMetTYrGlnAsnLeuGlnAspGlyLeu 100  
 Db 389 TCCAAGAACCTAGAAAAAGCATCAATATGTATCAAAACCTGACAGCAAGATGGGCTG 448  
 Qy 101 GluIysValHisLeuGluProValArgIleProHisTProGluArgGlyGlnSerIle 120  
 Db 449 GAGAAAGTTCACTGGAGCCAGTGAATACCCCACTGGAGAGGGAGAAATCAGCT 508  
 Qy 121 ValMetLeuGluProArgIleHisIleValAlaIleLeuGlyLeuGlySerSerIleGly 140  
 Db 509 GTGATGCTGGAGCCAGAAATTCATAGATACCATCTCGGCTTTGGCAGCAGCATTTGG 568  
 Qy 141 ThrProGluGlyIleThrAlaGlyValLeuValValThrSerPheAspGluLeuGln 160  
 Db 569 ACTCCTCCAGAGGATTACAGAGAGTTCGGTGGTGAACCTTTGATGAACTGCAAG 628  
 Qy 161 ArgArgAlaSerGluAlaArgGlyValIleValValValTYrAsnGlnProTYrIleAsnTYr 180  
 Db 629 AGAAGGGCTCAGAGAGCAAGAGGAGAGTGTGTGTATTAACCAACTTACATCAACTAC 688  
 Qy 181 SerArgThrValGlnTYrArgThrGlnGlyValAlaValAlaValAlaValAlaValAlaLeu 200  
 Db 689 TCAAGAGCGGTGCATACCAAGCCAGGGGGCGGTGGAGCTGCCAGAGTTGGGCTTTG 748  
 Qy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTYrSerProHisIleThrGlyIleGln 220  
 Db 749 GCATCTCTCATTCATCCGTGGCTCTCTTCATCTCAAGTCCACACAGGATTCAG 808  
 Qy 221 GluTYrGlnAspGlyValProIleIleProThrAlaCysIleThrValGluAspAlaGlu 240  
 Db 809 GAATACCAAGATGGGTGCCCAAGATTCCAACAGCTTATTAACGTGAGAGATGCCAGA 868  
 Qy 241 MetMetSerArgMetAlaSerHisGlyIleIleValIleGlnLeuIleMetGlyVal 260  
 Db 869 ATGATGTCAAAATGGCTTCTCATGGGATCAAAATGTCATTCACCTTAAAGATGGGCA 928  
 Qy 261 LysThrTYrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIleTYr 280  
 Db 929 AAGACCTACCAAGATACGATTCCTTCAACACTGTAGCAGAGATCATCTGGGAGCAAAATAT 988  
 Qy 281 ProGluGlnValValLeuValSerGlyHisIleuAspSerTTPAspValGlyGlnGlyVal 300  
 Db 989 CCAAGAACGGTGTGACTGTGTAGTGCATCTGGACACTGGAGTGGGAGGAGGTGCC 1048  
 Qy 301 MetAspAspGlyValGlyValAlaPheIleSerTTPGluAlaLeuSerLeuIleIleAspLeu 320  
 Db 1049 ATGGATGATGGGGTGGAGCCCTTATATCATGGAGGACCTCTCATATTATTAACATCTT 1108  
 Qy 321 GlyLeuArgProLYsArgTYrThrLeuArgLeuValLeuTYrThrAlaGluGlnGlnGly 340  
 Db 1109 GGGGCTGCTCCAAAGAGGACTCTGGGCTGGCTGTGCTGCTGCACTGCAGAAAGCAAGTGGGA 1168  
 Qy 341 ValGlyAlaPheGlnTYrTYrGlnLeuHisIleValAlaAsnIleSerAsnTYrSerLeuVal 360  
 Db 1169 GTTGTGTCTTCCAGTATTATCATAGTTACCAAGGTAATATTATTCACATACAGTCTGGTG 1228  
 Qy 361 MetGluSerAspAlaGlyThrPheLeuProHisArgIleGlnPheThrGlySerGlyIys 380  
 Db 1229 ATGGAGTCTGAGCGAGAACTCTTACCCATGGGCTGCAATTACGCGCAGTAAAG 1288  
 Qy 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400

Db 1289 GCCAGGCCATCATGAGAGGATTATGAGCTCTGCGAGGCCCTCAATATCATCAGTTC 1348  
 Qy 401 LeuSerHisGlyGluGlyTYrThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyVal 420  
 Db 1349 CTGAGCCATGAGAGAGGAGACAGACATCACTTTGGATCCAAAGCTGAGGTGGAGCC 1408  
 Qy 421 SerLeuLeuAspAspLeuTYrIleTYrPhePhePheHisHisSerHisGlyAspThrMet 440  
 Db 1409 AGTCTACTTGATGACTTATACAGATATTTCTTTCATCATCTCCACGGAGACACATG 1468  
 Qy 441 ThrValMetAspProLYsGlnMetAsnValAlaAlaValAlaValAlaValSerTYr 460  
 Db 1469 ACTGTCATGATCCAAAGCAGATGAATGTGCTGCTGCTGTTGGCTGTGTTCTTAT 1528  
 Qy 461 ValValAlaAspMetGluGluMetLeuProArgSer 472  
 Db 1529 GTTGTGACAGCATGAGAAATGCTCCTCAGGTCC 1564  
 RESULT 9  
 AA298139  
 ID AA298139 standard; cDNA; 1923 BP.  
 AC AA298139;  
 DT 11-MAY-2000 (first entry)  
 XX  
 DE Human signal peptide containing protein HSP-31 cDNA SEQ ID NO:165.  
 XX  
 KW Human; signal peptide-containing protein; HSP; diagnosis; cancer;  
 KW inflammation; cardiovascular disease; anticancer; anti-inflammatory;  
 KW antimicrobial; neuroprotective; cardioprotective; hepatocytic;  
 KW antidiabetic; gene therapy; cell proliferation; neurological disorder;  
 KW reproductive disorder; developmental disorder; arteriosclerosis;  
 KW cirrhosis; porphyria; acquired immune deficiency syndrome; anaemia;  
 KW asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;  
 KW Parkinson's disease; Huntington's diseases; ovulatory defect;  
 KW muscular dystrophy; ss.  
 KW  
 OS Homo sapiens.  
 XX  
 PN WO200000610-A2.  
 XX  
 PD 06-JAN-2000.  
 XX  
 PE 25-JUN-1999; 99WO-US44484.  
 XX  
 PR 26-JUN-1998; 98US-0090762.  
 PR 31-JUL-1998; 98US-0094983.  
 PR 01-OCT-1998; 98US-0102686.  
 PR 11-DEC-1998; 98US-0112129.  
 XX  
 PA (INCY-) INCYTE PHARM INC.  
 XX  
 PI Lal P, Tang YT, Gorgone GA, Corley NC, Guejler KJ, Baughn MR;  
 PI Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;  
 PI Bandman O;  
 XX  
 DR WPI; 2000-160673/14.  
 DR P-PSDB; AAY87254.  
 XX  
 PT New human signal peptide-containing proteins useful in treatment,  
 PT prevention and diagnosis of e.g. cancer, inflammation and  
 PT cardiovascular disease -  
 XX  
 PS Claim 9; Page 269-270; 327pp; English.  
 CC AA298109 to AA298242 encode AAY87224 to AAY87357 which represent the  
 CC human signal peptide-containing proteins HSP-1 to HSP-134. HSPs have  
 CC anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatocytic,  
 CC neuroprotective, cardioprotective, and antidiabetic activities, and can  
 CC be used in gene therapy. HSPs can be used to treat or prevent disorders  
 CC associated with decreased activity or function of HSP. Antagonists of  
 CC HSPs are used to treat or prevent disorders associated with increased

CC activity or function of HSP. Such diseases include cell proliferation  
(including cancer), inflammation, cardiovascular, neurological,  
CC reproductive or developmental disorders, (e.g. arteriosclerosis,  
CC cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,  
CC asthma, Crohn's disease, microbial or other infections, congestive or  
CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's  
CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSP  
CC nucleic acids can be used for the recombinant production of HSP, for  
CC detecting HSP in standard hybridisation and amplification assays (for  
CC diagnosis and monitoring), in gene therapy, as antisense, for  
CC triplex-forming or ribozyme therapeutics, for detecting related sequences  
CC or genetic variations, and for chromosomal mapping. HSP are also used to  
CC raise specific antibodies (Ab) and to screen for agonists and  
CC antagonists (potential therapeutic agents). Ab are used to diagnose, or  
CC monitor, HSP-related diseases (in usual immunoassays), as therapeutic  
CC antagonists, in competitive drug screens, and for purification of HSP  
CC from natural sources.

XX Sequence 1923 BP, 538 A; 439 C; 471 G; 475 T; 0 other;

## Alignment Scores:

Pred. No.:	2,21e-236	Length:	1923
Score:	2417.00	Matches:	472
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-09-745-763-36 (1-472) x AA298139 (1-1923)

QY 1 MetLysPheLeuIlePheAlaPhePheGlyValHISLeuLeuSerLeuCySerGly 20  
DB 179 ATGAATTCCTTATCTTCGATTTTCGGTGTTTCACTTTTATCCCTGCTGCTGG 238  
QY 21 LysAlaIleCyLeuAsnGlyIleSerLysArgThrPheGluGluIleLysGluIle 40  
DB 239 AAAGCTATATGCMAAATGGCATCTCTTAAGAGACTTTGAAAGAAATAAAGAAATA 298  
QY 41 AlaSerCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60  
DB 289 GCCACTGTGAGATGTTGCTTAAGCATCATCACTAGCTGTTTATGTTAAACCCAG 358  
QY 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
DB 359 AACGATCTCATAGACGATGGCACTTCGTGTATACGTGTGGACCCAGACTGAGTGGC 418  
QY 81 SerLysAsnLeuGluLysAlaIleGlnIleLeuTyrGlnAsnLeuGlnAspGlyLeu 100  
DB 419 TCCAAAGACCTTAGAAAAGCCATCCAAATTATGTATCCAAACCTCAGAGATGGCTG 478  
QY 101 GluLysValHISLeuGluProValArgTleProHISProGluArgGlyGluGluSerAla 120  
DB 479 GAGAAAGTTCACCTGAGCCAGTGAAGAAATACCCCACTGGGAGAGGGAGAAATCAGCT 538  
QY 121 ValMetLeuGluProArgGlyIleHISLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
DB 539 GTGATGCTGGAGCCAGAATTCATAGATAGCCATCCTGGGCTTGGCAGCAGCTTGG 598  
QY 141 ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160  
DB 599 ACTCTCCAGAGAGGATTAAGACAGAAAGTTCGTGTGTGACCTCTTTCATATACCTGAG 658  
QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValIleTyrAsnGlnProThrIleAsnTyr 180  
DB 659 AGAAGGGCTCAGAGAGAGAGAGAGATGTTGTTTAAACCACTTACATCACTAC 718  
QY 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValAlaLysValGlyValLeu 200  
DB 719 TCAAGCAGGCTGCAATACGAGACGAGGGGGGGTGGAGAGCTGCCAAGTTGGGCTTGG 778  
QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHISThrGlyIleGln 220  
DB 779 GCATCTCATTCATTCGATCCGTGGCTCTCTTCATCTACAGTCTCACACAGATATTGAG 838

QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
DB 839 GAATACCAAGATGGCGTGCCCAAGATTCCAAGACCTGTTATTCGGTGAAGATGCAAGAA 898  
QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyVala 260  
DB 899 ATGATGTCAAGAAATGGCTTCATTCATGAGATCAAAATTCATTCATTCAGTAAAGATGGGCA 958  
QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280  
DB 959 AAGACCTACCAAGATTCATTCCTTAAACACTGTACAGATCAGTGGAGCAAAATAT 1018  
QY 281 ProGluGlnValIleValLeuValSerGlyHISLeuAspSerTyrAspValGlyGlnGlyAla 300  
DB 1019 CCAAGACAGGTGTACTGCTGCTGAGACATCTGACAGCTGGATGTGGCAGAGGTGCC 1078  
QY 301 MetAspAspGlyGlyValAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320  
DB 1079 ATGATGATGGCGGTGAGCCTTATATCATGAGAACACTCTCATTTATTAAGATCTT 1138  
QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrProThrAlaGluGluGlnGly 340  
DB 1139 GGGCTGGCTCCAAAGAGAGACTGCGCTGCTGCTGAGCTGAGCAAGAAAGTGA 1198  
QY 341 ValGlyAlaPheGlnTyrTyrGlnLeuHISLysValAsnIleSerAsnTyrSerLeuVal 360  
DB 1199 GTTGCTGCTTCAGATTTATCATGTTACACAGGTAATATTTCCACTACAGCTGCTGG 1258  
QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380  
DB 1259 ATGAGATCTGACGAGAGAACCTTCTTACCCACTGGCTGCATATTCACCTGGCAAGAAAG 1318  
QY 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlyVal 400  
DB 1319 GCCAGGCGCATGAGAGAGAGGTATAGCTGTGAGAGCCCTCAATATCATCAGAGTTC 1378  
QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheThrIleGlnAlaGlyValProGlyAla 420  
DB 1379 CTGAGCATGAGAGAGAGAGAGCATCATCACTTTGGATCCAGCTGGAGTGCCTGAGACC 1438  
QY 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440  
DB 1439 AGTCTACTGATGACTTATACAAATATTTCTTCCATCATCTCCACGAGACACCATG 1498  
QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaValIleTPAlaValSerTyr 460  
DB 1499 ACTGTCAATGATCCAAAGCAGATGAATGTGCTGCTGCTGTGGCTGTGTTCTTAT 1558  
QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472  
DB 1559 GTTGTTCAGACATGAGAAATGCTGCTGATGCC 1594

RESULT 10  
ID ACCS0603 standard, cDNA; 2077 BP.  
XX ACCS0603;  
AC ACCS0603;  
DT 12-JUN-2003 (first entry)  
DE Human secreted protein coding sequence, SEQ ID 270.  
KW Cardiac; antiarrhythmic; antiarteriosclerotic; vasotropic; cyrostatic;  
KW vulnerability; antiinflammatory; nootropic; neuroprotective;  
KW antiparkinsonian; gene therapy; human; cardiovascular disorder;  
KW gene; ss.  
OS Homo sapiens.  
XX Homo sapiens.  
XX WO200295010-A2.  
XX 28-NOV-2002.  
XX PD

XX 19-MAR-2002; 2002WO-US09785.  
 PF 21-MAR-2001; 2001US-277340P.  
 PR 19-JUL-2001; 2001US-306171P.  
 PR 13-NOV-2001; 2001US-331287P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Rosen CA, Ruben SM;  
 PI WPI; 2003-129429/12.  
 DR  
 XX Novel human secreted proteins, useful for detecting, preventing,  
 PT diagnosing, prognosticating, treating and/or ameliorating  
 PT cardiovascular disorders such as arrhythmia -  
 XX  
 PS Claim 21; SEQ ID 270; 1881bp; English.  
 XX  
 CC The present invention relates to novel human secreted proteins  
 CC (ABR47633-ABR48145) and their coding sequences (ACC50344-ACC50856). The  
 CC proteins and their coding sequences are useful for the preparation of a  
 CC diagnostic or pharmaceutical composition for diagnosing or treating a  
 CC cardiovascular disorder (e.g., arrhythmia, tachycardia, cardiac arrest,  
 CC coronary arteriosclerosis and myocardial ischemia), neural disorders,  
 CC immune system disorders, muscular disorders, reproductive disorders,  
 CC gastrointestinal disorders, pulmonary disorders, renal disorders,  
 CC proliferative disorders and/or cancerous diseases and conditions, for  
 CC wound healing and epithelial cell proliferation, to treat inflammation or  
 CC infection, for treating thrombosis and arteriosclerosis, for treating or  
 CC preventing neural damage which occurs in neuronal disorders or  
 CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's  
 CC disease, to enhance bone and periodontal regeneration and aid in tissue  
 CC transplants or bone grafts, to prevent skin aging or hair loss, to  
 CC stimulate growth and differentiation of hematopoietic cells and bone  
 CC marrow cells when used in combination with other cytokines, to maintain  
 CC organs before transplantation or for supporting cell culture of primary  
 CC tissues, to increase or decrease differentiation or proliferation of  
 CC embryonic stem cells, or to modulate mammalian characteristics or  
 CC metabolism.  
 CC Note: The sequence data for this patent was published in electronic  
 CC format and is available from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 2077 BP; 597 A; 436 C; 482 G; 562 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2,46e-236 Length: 2077  
 Score: 2417.00 Matches: 472  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 25 Gaps: 0  
 US-09-745-763-36 (1-472) x ACC50603 (1-2077)  
 QY 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCySerGly 20  
 DB 132 ATGAAATTCCTTATCTTGCACTTTTCGGTGTGTCTCACTTTATCCCTGTGCTGGG 191  
 QY 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40  
 DB 192 AAACCTATATGCAAGATGCACTCTCTAAGAGCATTTTGAAGAAATTAAGAAATTA 251  
 QY 41 AlaSerCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValIYrGlyLysAlaGln 60  
 DB 252 GCCAGCTGTGAGATGTTGTCTTAAGCAATCACTAGCTGTTATGTGTAAGCCAG 311  
 QY 61 AsnArgSerTYrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
 DB 312 AACGATCTCTATGAGCGATGCGACTTGTGTGTACTGTGTGACCCAGACTGATGCGC 371  
 QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTYrGlnAsnLeuGlnGlnAspGlyLeu 100

DB 372 TCCAAGAACTAGAAAAAGCCATCAATTTGTACCAAAACCTGCAGCAAGATGGGCTG 431  
 QY GluLysValHisLeuGluProValArgIleProHisTrpPgluArgGlyGluSerAla 120  
 DB 432 GAGAAAGTTCACCTGTGAGCCAGTGAATATCCCATCGGAGAGGGGAGAAATCAAGCT 491  
 QY ValMetLeuGluProArgIleHisIleLysIleAlaIleLeuGlyLysSerSerIleGly 140  
 DB 492 GTGATGCTGAGCCAAAGATTCATTAAGATAGCCATCTCGGATCTTGGCAGCAATGGG 551  
 QY 141 ThrProGluGlyIleThrAlaGluValLeuValIleThrSerPheAspGluLeuGln 160  
 DB 552 ACTCCTCCAGAGAGCATTAACGAGAAAGTTCTGTGTGTGACTCTTTCGATGAATCGAG 611  
 QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValValIYrAsnGlnProTYrIleAsnTYr 180  
 DB 612 AGAAGGGGCTCAGAGCAAGAGGAGAGATTTGTTTATTAACCAACCTTACATCAACTAC 671  
 QY 181 SerArgThrValGluTYrArgThrGlnGlyValAlaGluAlaAlaLysValGlyAlaLeu 200  
 DB 672 TCAAGGACGTGCATTAACCAAGCGAGGGGCGGTGAGAGCTGCAGAGTTGGGCTTGG 731  
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTYrSerProHisTrpGlyIleGln 220  
 DB 732 GCATCTGCTCATTCGATCGGTGGCTCTCTTCATCACTACAGTCTTACACAGTATTCAG 791  
 QY 221 GluTYrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
 DB 792 GAATACCAAGATGTGGCGCCCAAGATTCCAACACCTGTATTACGGTGAAGATGCAGAA 851  
 QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyVala 260  
 DB 852 ATGATGTCAGAAATGGCTTCTCATGGAGTCAAAATGTGATTCAGCTTAAGATGGGGCA 911  
 QY 261 LysThrTYrProAspThrAspSerPheAsnThrValAlaGluIleTrpGlySerLysTYr 280  
 DB 912 AAGACTACCAAGATGATGATCTCTTCAACCTGACAGATCACTCGGAGCAATAT 971  
 QY 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnVala 300  
 DB 972 CCAGAACAGGTGTATCTGCTGACGTGACATCTGACAGCTGGAGTGTGGCAGGGTGC 1031  
 QY 301 MetAspAspGlyGlyValAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320  
 DB 1032 ATGATATATGGCGGTGAGGCTTTATATCATGGAGACCTGACTTATTAAGATCTT 1091  
 QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTrpThrAlaGluGluGlnGly 340  
 DB 1092 GCGCTGGCTCCAAAGAGAGACTCTGCGGCTGTGCTGTGAGCTGAGAAACAGGTGCA 1151  
 QY 341 ValGlyAlaPheGlnTYrTYrGlnLeuHisLysValAsnIleSerAsnTYrSerLeuVal 360  
 DB 1152 GTTGTCCTCTCCAGTATTAAGTATACAGTACACAGGTAAATATTTCACACTACAGCTGGTG 1211  
 QY 361 MetGluSerAspAlaGlyTYrThrPheLeuProThrGlyLeuGlnPheTrpGlySerGlyLys 380  
 DB 1212 ATGAGTCTGACGAGGAACCTCTTACCACTGGGCTGCACATTCACCTGGCAGTGAAG 1271  
 QY 381 AlaArgAlaIleMetGluGluValAlaMetSerLeuLeuGlnProLeuAsnIleThrGlyVal 400  
 DB 1272 GCCAGGCGCATCATGAGAGAGGTTATAGCTGTGCAAGCCCTCATATATCATCAGGTC 1331  
 QY 401 LeuSerHisGlyGlyTYrAspIleAsnPheTrpIleGlnAlaGlyValProGlyVala 420  
 DB 1332 CTGAGCCATGAGAGAGGAGACGACATCACTTTGTGATCCAGGTGAGTGGCTGGAGCC 1391  
 QY 421 SerLeuLeuAspAspLeuTYrTYrLysTYrPhePheHisHisSerHisGlyAspThrMet 440  
 DB 1392 AGTCTACTGATGACTTATACAGATATTTCTTTCATCATCACTCCACGAGACCATATG 1451  
 QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaIleValTrpAlaValIleSerTYr 460

Db 1452 ACTGTCATGATCCAAAGACATGATGTCGCTGCTGTTGGCTGTTCTTAT 1511  
 Qy 461 ValValAlaAapMecGlUglUmetLeuProArgSer 472  
 Db 1512 GTTGTTCAGACATGAGAAATGCTGCTAGTCC 1547  
 RESULT 11  
 AB271331  
 ID AB271331 standard; cDNA; 2077 BP.  
 AC AB271331;  
 XX 04-APR-2003 (first entry)  
 DT  
 DE Secreted protein-encoding gene 142 cDNA clone HRACT35, SEQ ID NO:152.  
 XX  
 KW Human; secreted protein; digestive disorder; gastrointestinal disorder;  
 KW mouth; oesophagus; stomach; small intestine; large intestine; liver;  
 KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;  
 KW immune disorder; inflammation; infection; wound healing; drug screening;  
 KW chromosome identification; chromosome mapping; cytostatic; gene therapy;  
 KW antiinflammatory; immunosuppressive; vulnereary; chromosome 8q22.2;  
 KW gene; ss.  
 KW  
 XX Homo sapiens.  
 OS  
 XX MO200276488-A1.  
 PN  
 XX 03-OCT-2002.  
 PD  
 XX 19-MAR-2002; 2002WO-US08276.  
 PF  
 XX 21-MAR-2001; 2001US-277340P.  
 PR 19-JUL-2001; 2001US-306171P.  
 XX 13-NOV-2001; 2001US-331287P.  
 PR  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX  
 PI Rosen CA, Ruben SM;  
 XX  
 DR WPI: 2003-029900/02.  
 DR P-PSDB; ABR00152.  
 XX  
 PT New human secreted proteins and nucleic acids, useful for detecting,  
 PT preventing, diagnosing, prognosticating, treating and/or ameliorating  
 PT e.g. gastrointestinal diseases and disorders, or cancers -  
 XX  
 XX Claim 21; Page 841-842; 1216bp; English.  
 XX  
 CC AB271150-AB271478 represent cDNAs corresponding to 178 human secreted  
 CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.  
 CC AB271479-AB271540 represent human secreted protein genomic fragments. The  
 CC invention also encompasses antibodies specific for the secreted proteins,  
 CC the use of the secreted proteins in drug screening, and recombinant  
 CC vectors and host cells comprising a nucleic acid of the invention. The  
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody  
 CC fragments specific for the secreted proteins, and modulators of protein  
 CC activity are useful for diagnosing, treating, ameliorating or preventing  
 CC digestive disorders. Such conditions include disorders of the mouth,  
 CC oesophagus, stomach, small intestine, large intestine, liver, biliary  
 CC tract and pancreas, and include cancers of these organs and tissues. The  
 CC secreted proteins and their nucleic acids may also be used in the  
 CC treatment of immune disorders, inflammation, infection,  
 CC hyperproliferative disorders, and to promote wound healing. Nucleic acids  
 CC of the invention may be used for chromosome identification, chromosome  
 CC mapping, in gene therapy, for identifying individuals from minute  
 CC biological samples, as hybridisation probes, and as molecular weight  
 CC markers. The present sequence represents a human secreted protein-  
 CC encoding cDNA clone of the invention.  
 XX  
 SQ Sequence 2077 BP; 597 A; 436 C; 482 G; 562 T; 0 other;

Alignment Scores:

Pred. No.: 2,466-236 Length: 2077  
 Score: 2417.00 Matches: 472  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Gaps: 0  
 Query Match: 100.00% Gaps: 0  
 DB: 25  
 US-09-745-763-36 (1-472) x AB271331 (1-2077)  
 Qy 1 MetLysPheLeuLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCysSerGly 20  
 Db 132 ATGAAATTCCTTATCTTCGCGATTTTTCGGTGCTTCACTTTATCCCTGCTCGG 191  
 Qy 21 LysAlaIleCysLysAaNGlyIleSerLysArgThrPheGluGluIleLysGluGlu 40  
 Db 192 AAACCTATATCAAGAAATGGCATCTCTAAGAGCATTTTGAAGAAATTAAGAAATA 251  
 Qy 41 AlaSerCysGlyAaPValAlaLysAlaIleIleAaAaLeuAlaValTyrGlyLysAlaGln 60  
 Db 252 GCCAGCTGTGGAGATGTTGGCTAAAGCAATCATCAACCTAGCTGTTATGTTAAAGCCGAG 311  
 Qy 61 AaAaArgSerTyrGluAaGleuAlaLeuLeuValAaPThrValGlyProAaGleuSerGly 80  
 Db 312 AACAGATCCTATAGAGCATTTGGCATTCTGCTTGATCTGTTGAGACCAAGCTAGTGGC 371  
 Qy 81 SerLysAaAaLeuGluLysAlaIleGlnIleMetTyrGlnAaAaLeuGlnAaPValLeu 100  
 Db 372 TCAGAAACCTTAGAAAGAGCCATCCAAATATATGTAACAAACCTTGACAGCAAGATGGCTG 431  
 Qy 101 GluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluGluSerAla 120  
 Db 432 GAGAAAGTTCACTGAGAGCCAGTGAGAAATCCCACTGGAGAGGAGGAAGATCAAGCT 491  
 Qy 121 ValMetLeuGluProAaGlyIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
 Db 492 GTGATGCTGAGCCAAATTCATAGATAGCCATCTCGGCTCTTGACAGCAGCATTTGG 551  
 Qy 141 ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAaPValLeuGln 160  
 Db 552 ACTCTTCAGAAAGCATTTACAGCAAGAGTTCTGAGTGTACCTTTTCAGTAACCTGAG 611  
 Qy 161 ArgArgAlaSerGluAaArgGlyLysIleValValTyrAaGlnProTyrIleAaAaTyr 180  
 Db 612 AGAAGGCGCTCAGAAAGCAAGAGGAGAAATGTTGTTATTAACCAACCTTACATCACTAC 671  
 Qy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaAaLysValGlyAlaLeu 200  
 Db 672 TCAAGGACGGTGCAATACCGAAACGAGGGCGGTGAGAGCTGCCAAGTTGGGGCTTTG 731  
 Qy 201 AlaSerLeuIleAaArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220  
 Db 732 GCATCTCTCATTTCAATCCGAGGCTCTCTTCCATCTACAGTCTTCCACAGAGATTCAG 791  
 Qy 221 GluTyrGlnAaPValProLysIleProThrAlaCysIleThrValGluAaPValGlu 240  
 Db 792 GAATACCAAGATGGCGGCCCAAGATTCCAACAGCTGATTAAGGGAAGATGCAAGAA 851  
 Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
 Db 852 ATGATGTCAGAAAGGCTTTCATGAGGATCAAAATTCATTCAGCTAAAGATGGGGCA 911  
 Qy 261 LysThrTyrProAaPThrAaPThrAaPThrValAlaGluIleThrGlySerLysTyr 280  
 Db 912 AAGACCTAACCAAGATGATCTTCTTCAACAGCTGTACAGAGATCATCGAGCAAAATAT 971  
 Qy 281 ProGluGlnValValLeuValSerGlyHisLeuAaPThrAaPValGlyGlnGlyAla 300  
 Db 972 CCAAGAACAGTTGATGCTGATGAGACATCTGACAGCTGGAGATGTTGGCAGAGGCTCC 1031  
 Qy 301 MetAaPhePglGlyGlyAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAaPLeu 320  
 Db 1032 ATGATGATGGCGGTGAGGCTTTATATCATGGAACACCTCTCATTAATAAGATCTT 1091

QY 321 GlyLeuArgProIleValGlnThrLeuArgLeuValLeuThrPheAlaGluGlnGlyGly 340  
DB 1092 GGGCTGGCTCCAAAGAGACTCTCGCGCTGGCTGCACTGCAGAAACAGGTGGA 1151  
QY 341 ValGlyAlaPheGlnIleThrGlnLeuHisValAsnIleSerAsnIleThrLeuVal 360  
DB 1152 GTGGTCCCTTCAGTATTAACAGTTCACAAAGGTAATATTTCCACTACAGTCTGGTG 1211  
QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLys 380  
DB 1212 ATGAGAGCTGACGAGAGAACTTCTTACCCACTGGCTGCAGATTCACTGGAGAGAAAG 1271  
QY 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400  
DB 1272 GCCAGGCCCATCAGAGAGAGTATGAGCTGCTGCAGGCCCTCAATATCATCTCAGGTC 1331  
QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheThrIleGlnAlaGlyValProGlyVal 420  
DB 1332 CTGAGCCATGAGAGAGGACAGACATCAACTTTGGATCCAAAGCTGAGTCCGAGGCC 1391  
QY 421 SerLeuLeuAspAspLeuIleThrPhePhePheHisSerHisGlyAspThrMet 440  
DB 1392 AGTCTACTGATGACTTATACAGATATTTCTTCCATCACTCCACGAGAGACCATG 1451  
QY 441 ThrValMetAspProIleGlnMetAsnValAlaAlaValAlaValSerIle 460  
DB 1452 ACTGTCATGATCCAAAGACAGATGATGTTCTGCTGCTGTTGGCTGTTGTTCTTAT 1511  
QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472  
DB 1512 GTTGTTCAGACATGAGAAATGCTGCTAGGTC 1547

## RESULT 12

AAA40493 standard; cDNA; 1767 BP.

AAA40493;

16-NOV-2000 (first entry)

Human fetal kidney cDNA fragment AM282\_11.

Secreting protein; cytosolic; immunostimulatory; antimicrobial;  
antiviral; immunosuppressive; antiinflammatory; vulnary; cytokine;  
cell proliferation; differentiation; regulator; treatment; tumor;  
autoimmune disease; inflammatory disorder; wound; microbial infection;  
viral disease; graft versus host reaction suppression; ss.

Homo sapiens.

WO200037630-A1.

29-JUN-2000.

22-DEC-1999; 99WO-US31005.

23-DEC-1998; 98US-0220876.

(GENY) GENETICS INST INC.

Jacobs K, McCoy JM, Lavallie ER, Collins-Racie, LA, Evans C,

Merberg D, Treacy M, Bowman MR;

WPI; 2000-442661/38.

P-PSDB; AAB10229.

Secreted human proteins AS296-11 and AS34-11, useful for treating

tumors, autoimmune diseases, inflammatory disorders, wounds, microbial

infections and viral diseases -

CC cytosolic, immunostimulatory, antimicrobial, antiviral,  
CC immunosuppressive, antiinflammatory and vulnary activity and which act  
CC as cytokine, cell proliferation or differentiation regulators. (1)  
CC is useful for treating tumors, autoimmune diseases, inflammatory  
CC disorders, wounds, microbial infections and viral diseases. (1) is also  
CC useful for suppressing graft versus host reaction. AAA40490-AA40580  
CC represent cDNA fragments that encode the secreted proteins  
CC AAB10226-B10288 described in the method of the invention.

XX Sequence 1767 BP; 512 A; 384 C; 418 G; 453 T; 0 other;

## Alignment Scores:

Pred. No.:	Length:
Score:	2410.00
Percent Similarity:	99.79%
Best Local Similarity:	99.79%
Query Match:	99.71%
DB:	21
Gaps:	0

US-09-745-763-36 (1-472) x AAA40493 (1-1767)

QY 1 MetLeuPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCySerGly 20  
DB 6 ATGAAATTCCTTATCTTGGCATTTTCGAGGTTCACCTTTATCCCTGCTGCGG 65  
QY 21 LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluIle 40  
DB 66 AAGCTATATGCAAGATGACATCTTACAGAGACTTTTGAAGAAATTAAGAAATA 125  
QY 41 AlSerCysGlyAspValAlaValAlaIleIleAsnLeuAlaValIleGlyLysAlaGln 60  
DB 126 GCCAGCTGTGAGATGTTGCTTAAAGCATCATCACTAGTATTGTTAAAGCCAG 185  
QY 61 AsnArgSerIleGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
DB 186 AACAGATCTTATGAGCGATTGGCACTTGTGATTACTGTGAGCCCACTGAGTGC 245  
QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetIleThrGlnAsnLeuGlnIleAspGlyLeu 100  
DB 246 TCCAGAACTTAAAGAAAGCAATCCAAATATGTAACAAACCTGCAGAGATGGCTG 305  
QY 101 GluLysValHisLeuGluProValArgIleProHisIleProGluArgGlyGluSerAla 120  
DB 306 GAGAAATTCACCTGAGCGCAGAGATATCCCACTGAGAGAGGAGAAATCATGCT 365  
QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerIleGly 140  
DB 366 GTGATGCTGAGCGCAAGAAATTCATTAAGATGACATCTGAGTCTGGCAGCAGATTGGG 425  
QY 141 ThrProGluGluIleThrAlaGluValLeuValIleThrSerPheAspGluLeuGln 160  
DB 426 ACTCTTCAGAAAGCATTTACAGAGAGTCTGTGTGAGCTCTTTCAGAGAACTGAG 485  
QY 161 ArgArgLysSerGluAlaArgGlyLysIleValIleValIleValIleValIleValIle 180  
DB 486 AGAAGGCTTCAGAGAGAGAGAGAGATGTTGTTTAAACCACTTATCATCACTAC 545  
QY 181 SerArgThrValGlnIleThrArgGlnGlyValIleValIleValIleValIleValIle 200  
DB 546 TCAAGAGCGGTGCAATACCAAGACGAGGCGGTGAGAACTCCCAAGGCGGCTTGG 605  
QY 201 AlSerLeuIleArgSerValAlaSerPheSerIleThrProHisIleThrGlyIleGln 220  
DB 606 GCACTCTCATGATCGATCGTGGCTCTCTTCCATCTACAGTCTCCACAGAGATTACG 665  
QY 221 GluIleGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGln 240  
DB 666 GAATACAGAGATGGCGGCCAGATTCCAAAGCGCTATTTACGGTGGAGAAATGAGAA 725  
QY 241 MetLeuSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
DB 726 ATGATGCAAGAAATGCTCTCATGAGATCAAAATGTCATTCAGTAAAGATGGGGCA 785



```

Oy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluLeuThrGlySerLysTyr 280
Db 786 AAGACCTCCAGATGATGATTCCTTCAACACTGAGAGATGATCACTGGAGCAAAAT 845
Oy 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnVal 300
Db 846 CCAGAACAGGTGTGACTGCTGATGATCATCTGACACTGGAGATGTTGGCAGGGTGC 905
Oy 301 MetAspAspGlyGlyGlyValAlaPheLeuSerTrpGlnAlaLeuSerLeuLeuLeu 320
Db 906 ATGATGATGAGCGGAGGACCTTATATCATGGAGAACCTCTCACTTAAATTAATTT 965
Oy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTrpThrAlaGlyGlnGlyGly 340
Db 966 GGGCTGCTCCAAAGAGACTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1025
Oy 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAlaLeuSerAsnTyrSerLeu 360
Db 1026 GTTGGTGCCTTCCAGTATATATCATGATCAAGTAATTAATTTCCACTACAGTCTGG 1085
Oy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGly 380
Db 1086 ATGAGTCTGACGCGAGAACCTTCTTACCCTGAGCTGCAATTCACCTGGAGTAAG 1145
Oy 381 AlaArgAlaLeuMetGluGlnValMetSerLeuGlnProLeuAsnLeuThrGlnVal 400
Db 1146 GCCAGGCGCATCTATGAGAGAGGTTATGAGCTGCTGCGAGCCCTCAATATCATCAG 1205
Oy 401 LeuSerHisGlyGlyGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGly 420
Db 1206 CTGAGCCATGAGAGAGGACAGACATCAACTTTTGATTCAGAGCTGGAGTCTGAGAC 1265
Oy 421 SerLeuLeuAspAspLeuTyrTyrPhePhePheHisSerHisGlyAspPheMet 440
Db 1266 AGCTACTGTGATGATCTTATACAGATATTTCTTCCATCACTCCACGAGACCCATG 1325
Oy 441 ThrValMetAspProLysGlnMetAsnValAlaAlaValAlaValAlaValSer 460
Db 1326 ACTGTCATGATGCCAAAGCATGATGATGTTGCTGCTGCTGTTGGGCTGTGTTCTT 1385
Oy 461 ValValAlaAspMetGlyGlnMetLeuProArgSer 472
Db 1386 GTTGTTCAGACATGAGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1421

```

RESULT 13  
AAV02296  
ID AAV02296 standard; DNA; 1778 BP.  
XX  
AC AAV02296;  
XX  
DT 21-MAY-1998 (first entry)  
XX  
DE Human secreted protein AM282 full-length cDNA clone.  
XX  
KW Secreted protein; AM282; cytokine; human; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 17..1435  
FT sig\_peptide /tag= a  
FT mat\_peptide /tag= b  
FT /tag= 89..1432  
FT /tag= b  
XX  
PN MO9739030-A2.  
PD 23-OCT-1997.  
XX  
PF 16-APR-1997; 97WO-US06475.  
XX  
PR 13-JAN-1997; 97US-0783520.

```

PR 18-APR-1996; 96US-0634325.
XX  
XX (GENM) GENETICS INST INC.  
XX  
P1 Jacobs K, Lavallie ER, McCoy JM, Werberg D, Racie LA;  
P1 Spaulding V;  
XX  
XX MPI: 1997-526400/48.  
DR P-PSDB: AAM33604.  
XX  
XX New isolated secretory proteins AM340, AM282 and AK583 - possibly  
PT have cytokine, cell proliferation/differentiation regulating,  
PT immunomodulating activities, etc.  
XX  
XX Claim 15; Page 44-45; 59pp; English.  
XX  
XX This cDNA clone encodes a protein (see W33604) designated AM282.  
CC It was identified as "Yf95b10.r1 human EST 30142.5" (Genbank  
CC accession No. R77830) in a database search using a partial AM282  
CC clone (see T97398) obtained from a human foetal kidney cDNA  
CC library using methods selective for cDNAs encoding secreted  
CC proteins. AM282 is deposited in ATCC 98026 together with clones  
CC AM340 (see T97397) and AK583 (see V02297), which are also claimed.  
CC AM282 protein can be used in a claimed method for preventing,  
CC treating or ameliorating a medical condition. It may exhibit  
CC cytokine, cell proliferation (either inducing or inhibiting) or  
CC cell differentiation (either inducing or inhibiting) activity or  
CC may induce production of other cytokines in certain cell  
CC populations. It may also exhibit e.g. immune stimulating or  
CC suppressing activity, haematopoiesis regulating activity, tissue  
CC growth activity, activin/inhibin activity, chemocactic or  
CC chemokine activity, haemostatic or thrombolytic activity,  
CC receptor/ligand activity, anti-inflammatory activity, tumour  
CC inhibition activity, or other activities. No evidence of any of  
CC these activities is given in the specification.  
XX  
SQ Sequence 1778 BP; 514 A; 386 C; 422 G; 456 T; 0 other;  
SQ  
Alignment Scores:  
Pred. No.: 1,02e-235 Length: 1778  
Score: 2410.00 Matches: 471  
Percent Similarity: 99.79% Conservative: 0  
Best Local Similarity: 99.79% Mismatches: 1  
Query Match: 99.71% Indels: 0  
DB: 18 Gaps: 0  
US-09-745-763-36 (1-472) x AAV02296 (1-1778)  
Oy 1 MetLysPheLeuIlePheAlaPheGlyGlyValHisLeuSerLeuCySerGly 20  
Db 17 ATGAAATTCCTTATCTTCCGATTTTTCGGTGGTGTTCACCTTTATCCGTGTCTGGG 76  
Oy 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGlnIleLeuGlnIle 40  
Db 77 AAACCTATATGCAAGATGCGATCTCAAGGACCTTTGAAGAAATTAAGAAAGATA 136  
Oy 41 AlaSerCyGlyAspValAlaValAlaIleLeuSerLeuValTyrGlyValAlaGln 60  
Db 137 GCCAGCTGTGAGATGTTGCTAAAGCAATCATCACTAGCTGTTATGTAAGGCCAG 196  
Oy 61 AsnArgSerTyrGluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80  
Db 197 AACGATCTTATGCGCATTTGGCACTTCTGTTGATATCTGTGAGCCAGACTGAGTGC 256  
Oy 81 SerLysAsnLeuGlnLysAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100  
Db 257 TCCAAAGAACCTAGAAAAGCCATCCCAATTATGTAACCAAACTGACGAAAGTGGCTG 316  
Oy 101 GlnLysValHisLeuGlnProValArgIleProHisTrpGluArgGlyGlnGlnSer 120  
Db 317 GAGAAAGTTCACTCGGAGCCAGTGAATATCCCACTGGAGAGGAGAGAAATTCAGCT 376  
Oy 121 ValMetLeuGlnProArgIleHisIleValAlaIleLeuGlyLeuGlySerSerIleGly 140

```



```

Db      377 GTGATGCTGGAGCCAAAGATTTCATAGATAGCATCCGCGCTTGGCAGCAGCATTTGGG 436
Qy      141 ThrProGluGlyIleThrAlaGluValIleuValIleuThrSerPheAspGluLeuGln 160
Db      437 ACTCTCCAGAAAGCATTTACACAGAGATCTTGCGTGACCTCTTTCAGATGAACTGCAG 496
Qy      161 ArgArgAlaSerGluAlaArgGlyIleValIleValIleValIleValIleValIleVal 180
Db      497 AGAAGGGCTCCTCAGAAAGCAGAGAGAGATTTGTTTATACCAACCTTACATCAGCTAC 556
Qy      181 SerArgThrValGlnTyrArgThrGlnGlnValAlaGluAlaAlaValValIleVal 200
Db      557 TCAAGGACGGTGACATACCGAAGCGAGGGGGGAGGAGCTGCCAAGGTGGGGGCTTTG 616
Qy      201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
Db      617 GCATCTCTCATTCGATCCGTCGCTCTTTCATCTACAGTCTTCACACAGGATTTGAG 676
Qy      221 GluTyrGlnAspGlyValIleProIleIleProThrAlaCysIleThrValIleGluAspAlaGlu 240
Db      677 GAATACCGAGATGGGCTGCCAAGATTCACACAGCTGTATTTACGGTGGAAAGATGACAA 736
Qy      241 MetMetSerArgMetAlaSerHisGlyIleIleValIleValIleGlnLeuIleMetGlyAla 260
Db      737 ATGATGTCAAGAAATGGCTTTCATGAGATCAAAATTTGATTCACATCAAGCTAAAGATGGGGCA 796
Qy      261 LysThrTyrProAspThrAspSerPheLeuThrValAlaGluIleThrGlySerIleTyr 280
Db      797 AAGACCTAACCAAGATACGATTCCTTCAACACTGTAGAGAGATCAGCGGAGCAAAAT 856
Qy      281 ProGluGlnValIleValIleValIleSerGlyIleLeuAspSerTyrAspValIleGlnIleVal 300
Db      857 CCAAGAACGATTTGATCTGTCAGTGCATCTGCAGCACTGCAGCTGGAGATTTGGGAGGGTCC 916
Qy      301 MetAspAspGlyGlyIleAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320
Db      917 ATGATGTAGTGGCGGGAGCGCTTATATCATGCGAAGCACTCTCATTTATAAATCTT 976
Qy      321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTyrThrAlaGluGlnGlnGly 340
Db      977 GGGCGCTCCAAAGAGACTCTGGCGCTGCTGCAGCTGCAGCGAGAACTAGTGGGA 1036
Qy      341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
Db      1037 GTTGGTGCCTTCCAGTATTTATCAGTTACCAAGGAAATATTTCCAACTACAGCTGCTG 1096
Qy      361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLys 380
Db      1097 ATGGAGCTCGACGAGCAACCTTTTACCACTGGGCTGCATTTCACTGGCGAGTAAAG 1156
Qy      381 AlaArgAlaIleMetGluGlnValIleMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db      1157 GCCAGGGCCATCATGAGAGAGGTTTGTAGCCCTGCAGCGCCCTCAATATCATCACTAGCTC 1216
Qy      401 LeuSerHisGlyGlnGlyThrAspIleAsnPheTyrIleGlnAlaGlyValIleProGlyAla 420
Db      1217 CTGAGCCATGAGAGAGGAGCAATCATCTTTGATTCAGAGCTGAGTGCCTGAGGCC 1276
Qy      421 SerLeuLeuAspAspLeuTyrTyrPhePhePheHisSerHisGlyAspThrMet 440
Db      1277 AGTCACTTGAATGATTTATACAGATATTTCTTCCATCACTCCACAGAGACACCATG 1336
Qy      441 ThrValMetAspProLysGlnMetCysValAlaAlaAlaValTyrValValIleSerTyr 460
Db      1337 ACTGCATGAGATCCAAACAGATGATGTGCTGCTGTTGGCTGTTGTTGTTTCTTAT 1396
Qy      461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db      1397 GTTGTTCAGACATGAAAGAAATGCTGCTAGGTC 1432

```

RESULT 14  
AAK94491

```

ID - AAK94491 standard; cDNA; 1860 BP.
XX
AC AAK94491;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA, SEQ ID NO: 3328.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI: 2001-524255/58.
DR P-PSDB: AAK93559.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 8; SEQ ID NO 3328; 1380bp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesized by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 1860 BP; 515 A; 423 C; 451 G; 471 T; 0 other;

```

Alignment Scores:

Pred. No.:	Length:	Score:
7,14e-235	1860	2402.00
Percent Similarity:	Matches:	469
99.79%	Conservative:	2
Best Local Similarity:	Mismatches:	1
99.38%	Indels:	0
Query Match:	Gaps:	0

US-09-745-763-36 (1-472) x AAK94491 (1-1860)

```

Qy      1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCysSerGly 20
Db      129 ATGAAATTCCTTATCTTCGCAATTTTTCGGTGGTGTTCACCTTTTATCCCTGTGCTGGG 188
Qy      21 LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40
Db      189 AAGCTATATCAAGATGCGATCTCTAAGAGGACTTTTGAAGAATAGAGAGAAATA 248
Qy      41 AlaSerCysGlyAspValAlaAlaValAlaIleLeuLeuAlaValTyrGlyLysAlaGln 60
Db      249 GCCAGCTGTGAGATGTTGCTTAAAGCAATCTCAACCTAGCTGTTTATGAGTAAAGCCAG 308
Qy      61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValIleGlyProArgLeuSerGly 80

```

```

Db      ||| 309 AACAGATCTTATGACGATTGGCAGCTTGGTTGATCTGTTGACCCAGACGATGAGTGGC 368
Oy      81 SerLySAsnLeuGluValAlaIleGlnIleMetYrGlnAsnLeuGlnIAspGlyLeu 100
Db      369 TCCAAAGACCTTGAAGAAAGCCATCCAAATATGATGACAAACCTCAGCAAAATGGGCTG 428
Oy      101 GluLySValHisLeuGluProValArgIleProHisSTTPGluArgGlyGluGluSerAla 120
Db      429 GAGAAAGCTTACCTTGAGCGCACTGAGAAATACCCCACTGGAGAGGGGAGAGAAATCACT 488
Oy      121 ValMetLeuGluProArgIleHisIleValIleAlaIleLeuGlyLeuGlySerIleGly 140
Db      489 GTGATGCTGAGACCAAGAAATTCATAGATAGCCATCTGGGCTTGGCAGCAGCATGGG 548
Oy      141 ThrProProGluGlyIleThrAlaGluValIleValIleThrSerPheAspGluLeuGln 160
Db      549 ACTCTCCAGAGGAGGATTAACAGCAAGTCTGGTGGTGAACCTCTTTCGATGAACTGAG 608
Oy      161 ArgATGAlaSerGluAlaArgGlyLysIleValIleValIleYrAsnGlnProTyrIleAsnTyr 180
Db      609 AGAAGGGCTTCAGAAAGCAAGAGGAGAGATTGTTGTTTATACCAACCTTACATCACTAC 668
Oy      181 SerArgThrValGlnIleTyrArgThrGlnGlyAlaValAlaGluAlaIleValIleValLeu 200
Db      669 TCAAGCAGGCTGCATATACGAAACGAGGGGGCGGTGGAAGCTGCCAAGGTTGGGCTTGG 728
Oy      201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisSThGlyIleGln 220
Db      729 GCATCTTCATTCGATCCGCTGGCTCCCTCTTCATCTCAAGCTCCACACAGATATTCAG 788
Oy      221 GluTyrGlnAspGlyValAlaProLysIleProThrAlaCysIleThrValIleGluAspAlaGlu 240
Db      789 GAATCCAGAGATGGCGTCCCAAGATTCACACGCTGATATTCGCTGGAATCAGACAA 848
Oy      241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db      849 ATGATGTCAGAAATGCTTCTCATGGGATCAAAATTTGTCATTCACATTAAGATGGGGCA 908
Oy      261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIleTyr 280
Db      909 AAGACCTTACCCAGATACGATTCCTTCAACACCTGAGCAGATATCAGGGGCAAAATAT 968
Oy      281 ProGluGlnValIleLeuValSerGlyHisLeuAspSerTTPAspValGlyGlnIleVala 300
Db      969 CCAAGAACAGTTTGTACTGCTGCTGACATCTGACAGCTGGAGATCTTGGGAGAGGTGCC 1028
Oy      301 MetAspAspGlyGlyIleAlaPheIleSerTTPGluAlaLeuSerLeuIleLysAspLeu 320
Db      1029 ATGATGATGGCGGTGGAGCCTTTATATCATGGAAGCACTCTCACTTATTAAGATCTT 1088
Oy      321 GlyLeuAspProLysArgThrLeuArgLeuValIleLeuTTPAlaGluGlnIleGly 340
Db      1089 GGGCTGCTCCAAAGAGACTCTGGCGCTGCTCTGACGACGAGAAACAAAGTGGGA 1148
Oy      341 ValGlyAlaPheGlnTyrTyrGlnLeuHisIleValIleAsnIleSerAsnTyrSerLeuVal 360
Db      1149 GTTGATGCTTCCAGTATTTATAGTTACACAGGTAATATTTCCACTACAGCTCTGGTG 1208
Oy      361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLys 380
Db      1209 ATGAGTCTGACGAGCAACCTTTTACCACCTGGGCTGCATTCACCTGAGGAGTAAAG 1268
Oy      381 AlaAlaGlyAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db      1269 GCCAGGGCCATCTGAGAGAGGTTTATGAGCCCTGCGACGCCCTCAATATCACTACAGCTC 1328
Oy      401 LeuSerHisGlyGlyGlyIleThrAspIleAsnPheTTPIleGlnAlaGlyValProGlyAla 420
Db      1329 CTGAGCCATGAGAGAGGACAGACATCACTTTTGATTCAGAGCTGAGAGTGGCTTGAAGCC 1388
Oy      421 SerLeuLeuAspAspLeuTyrIleTyrPhePhePheHisHisIleSerHisGlyValAspThrMet 440

```

```

Db      1389 AGCTACTGATGACTTATACAGTATTTCTTTCATCATCACTCCACGAGACACCATG 1448
Oy      441 ThrValMetAspProLysGlnMetAsnValAlaIleAlaValIleTTPAlaValIleSerTyr 460
Db      1449 ACTGTCATGATCCAAAGCAGATGAAATGTCTCTGCTGCTGTTGGCTGTTGTTCTTAT 1508
Oy      461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db      1509 GTTGTTGAGACATGAGAAAGAAATGCTGCTAGTGC 1544

RESULT 15
AAH99703 standard; cDNA; 1895 BP.
AC AAH99703;
DT 16-OCT-2001 (first entry)
DE Human protein encoding cDNA sequence SEQ ID NO:538.
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; vitruclide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnery; antileuk; osteopathic; eczema;
KW dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antiaphyllactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder; ss.

Homo sapiens.
OS
MO200153455-A2.
PN 26-JUL-2001.
PD
XX 22-DEC-2000; 2000WO-US35017.
PF
XX 23-DEC-1999; 99US-0471275.
PR 21-JAN-2000; 2000US-0481275.
PR 25-APR-2000; 2000US-0552317.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Dymnac RT;
XX WPI; 2001-457603/49.
XX P-PSDB; AAM25762.
XX
XX Isolated human polynucleotides encoding polypeptides, useful for the
XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
XX Claim 1; Page 591; 1217pp; English.
PS
XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
XX AAM25963. The proteins can have activities based on the tissues and
XX cells they are expressed in, such as: antiinflammatory; antirheumatic;
XX antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
XX central nervous system; vitruclide; anti-HIV; fungicide; antimutagen;
XX cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
XX antileuk; osteopathic; dermatological; antiallergic; antiaesthetic;
XX antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
XX antiparkinsonian; and immunostimulant. The proteins and polynucleotides
XX encoding them can be used in gene therapy, antisense therapy and vaccine
XX production. The proteins and polynucleotides are useful for screening for
XX agonists or antagonists of a protein and for the treatment and diagnosis
XX of disorders associated with the activity of a protein e.g. inflammation,
XX

```

CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
 CC neurological disorders.

XX Sequence 1895 BP: 530 A; 439 C; 450 G; 476 T; 0 other:

# Alignment Scores:

Pred. No.:	1,74e-228	Length:	1895
Score:	2339.50	Matches:	468
Percent Similarity:	98.32%	Conservative:	0
Best Local Similarity:	98.32%	Mismatches:	4
Query Match:	96.79%	Indels:	4
DB:	22	Gaps:	1

US-09-745-763-36 (1-472) x AAH9703 (1-1895)

```

Qy 1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGly 20
Db 126 ATGAAATTCCTTATCTTCGCAATTTTCGGTGTGTCACTTTATCCCTGTCTGGG 185
Qy 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40
Db 186 AAGACTATATCCAGAAATGCGATCTCTAAGAGCATTTTGAAGAAATTAAGAAATA 245
Qy 41 AlasercySglYaspValAlaLysAlaIleLeuLeuAlaValIYrgLYLysAlaGln 60
Db 246 GCCAGCTGTGAGATGTGTGTAAGCAATCATCACTAGCTGTATGTGTAAGCCAG 305
Qy 61 AsparGserTYrgLuarGleuAlaLeuLeuValAspThrValIGYrProArgLeuSerGly 80
Db 306 AACGATCTCTATGAGCCATGTCGACTCTGTGATACCTGTGAGCCAGCATAGTGC 365
Qy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTYrgLnaAsnLeuGlnAspGlyLeu 100
Db 366 TCCAGAACTCTAGAAAGCAAGCATTCATTTATGACCAAACTCTGACAGAAAGGCTG 425
Qy 101 GluLysValHisLeuGluProValArgIleProHisTrpGluArgGlyGluGluSerAla 120
Db 426 GAGAAAGTTCACCTGTGAGCCAGTGAATATACCCCACTGGAGAGGAGGAAGATTCAG 485
Qy 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
Db 486 GTGATGCTGGAGCCAGAAATTCATAGATACCAATCCCTGCTTGGCAGCAGCATTGG 545
Qy 141 ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160
Db 546 ACTCTCTCCAGAAAGCATTTACAGCAAGATCTGTGTGTGAGACTCTTTCGAGAACTG 605
Qy 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTYraGlnProTYrIleAsnTYr 180
Db 606 AGAAGGGGCTCAGAAAGCAAGAGGAGATTGTGTTATTAACCAACTTATCATCACTAC 665
Qy 181 SerArgThrValGlnTYrArgThrGlnGlyAlaValGluAlaIleLysValGlyAlaLeu 200
Db 666 TCAGAGACGGTGCATATCCAGAACCCAGGGGGGGTGAAGCTGCCAAGTGGGGCTTTG 725
Qy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTYrSerProHisIleThrGlyIleGln 220
Db 726 GCATCTCTCATTCATTCGTCGTCCTCTTCATCTACATGATCTCTCAGCAGATTTAG 785
Qy 221 GluTYrgLnaPrgLYValProLysIleProThrAlaCySileThrValGluAspAlaGlu 240
Db 786 GAATATCCAGATGGGTGCCAAGATTCCAACAGCTGTATTAAGGTGAAGATGCAGAA 845
Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db 846 ATGATGTCAGAAATGGCTTTCATAGGATCAAAATTTTCATTCAGCTAAAGATGGGGCA 905

```

```

Qy 261 LysThrTYrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerTYr 280
Db 906 AAGACCTACCCAGATACTGATTCCTTCAACACTGTAGCAGATTCACCTGGAGCAATAT 965
Qy 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300
Db 966 CCAGAACAGGTGTACTGCTGCTGAGTGCATCTGACAGCTGGAGTGTGGGAGGGTGC 1025
Qy 301 MetAspAspGlyGlyGlyAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320
Db 1026 ATGATGATGCGGTGAGCTTTATATCATGGAAGCACTTCATTTATTAAGATCTT 1085
Qy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTrpThrAlaGluGlnGlyGly 340
Db 1086 GGGGTGCTCCAAAGAGACTCTGCGCTGTCTCTGACCTGACAGAAACAAGTGA 1145
Qy 341 ValGlyAlaPheGlnTYrTYrgLnaLeuHisLysValAsnIleSerAntTYrSerLeuVal 360
Db 1146 GTTGTCCTTCCAGATTTATCAGTTACACAGGTAATATTTCCAACTACAGCTGTG 1205
Qy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
Db 1206 ATGAGTCTACCGAGAACCTTTATCCCACTGGGCTGCAATTCATGCGCAGGAAAG 1265
Qy 381 AlaArgAlaIleMetGluGluValMetSerLeuGlnProLeuAsnIleThrGlnVal 400
Db 1266 GCCAGGSCCATCATGAGAGGATTATGAGCTCTGAGCCCTCAATATCATCTAGTTC 1325
Qy 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420
Db 1326 CTGAGCCATGAGAAAGGAGCAGACATCACTTTGGATCCAGCTGAGAGTGGCTGAGCC 1385
Qy 421 SerLeuLeuAspLeuTYrLysTYrPhePhePheHisSerHisGlyAspThrMet 440
Db 1386 AGTCTACTATGACTTATACAGATATTTCTTCCATCATCTCCACGAGACACCATG 1445
Qy 441 ThrValMetAsp---ProLysGlnMetAsnVal-AlaAlaIleVal-TrpAlaValAla 459
Db 1446 ACTGTCAATGGATCCAAAGCAGATGATGTTTGGCTGCTGCTGTTTGGGCTGTGTTT 1505
Qy 459 exTYrVal-ValAlaAspMetGluLeuMetLeuProArgSer 472
Db 1506 CTATATGTGTGAGACATGAGAAATGCTGCTGATGCTC 1547

```

Search completed: December 22, 2003, 23:51:21  
 Job time : 398 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: December 22, 2003, 23:13:56 / Search time 2345 Seconds  
(without alignments)  
4891.988 Million cell updates/sec

Title: US-09-745-763-36  
Perfect score: 2417  
Sequence: 1 MKFLIFAFGFGVHLISLCSG.....AWAVSVYVADMEMLPRS 472

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 22781392 segs, 12152238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seg length: 0  
Maximum DB seg length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DRV=xlp  
-Q=cgnt2\_1/USFPO.spool/p/US9745763/runat.22122003\_113317\_25262/app\_query.fasta.1.647  
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=400 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US09745763 @CGN 1.1 2874 @runat.22122003\_113317\_25262 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOUTERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_hcl:\*  
10: gb\_est1:\*  
11: gb\_est2:\*  
12: gb\_hlc:\*  
13: gb\_est3:\*  
14: gb\_est4:\*  
15: gb\_est5:\*  
16: em\_estfun:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_dhg:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vr1:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2133	88.2	1739	11 AK032972	AK032972 Mus muscu
2	2133	88.2	1805	11 AK075686	AK075686 Mus muscu
3	2124	87.9	1868	11 BC010977	BC010977 Mus muscu
4	1633	67.6	1024	11 BC012019	BC012019 Homo sapi
5	1539	63.7	1050	13 BX439467	BX439467 BX439467
6	1431	59.2	1038	13 BX355940	BX355940 BX355940
7	1352	55.9	909	13 BQ878966	BQ878966 BX360507
8	1312.5	54.3	1126	13 BX360507	BX360507 BX360507
9	1301.5	53.8	902	10 BG751497	BG751497 602730076
10	1290	53.4	919	13 BX335995	BX335995 BX335995
11	1277	52.8	866	13 BU146905	BU146905 AGENCOURT
12	1253	51.8	739	14 CB958693	CB958693 AGENCOURT
13	1252	51.8	1012	13 BX416896	BX416896 BX416896
14	1244	51.5	996	13 BX355939	BX355939 BX355939
15	1213	50.2	926	13 BQ938234	BQ938234 AGENCOURT
16	1204	49.8	998	13 BX460463	BX460463 BX460463
17	1200	49.6	804	14 CD352906	CD352906 UI-M-GLO-
18	1190	49.2	3151	11 BC017373	BC017373 Mus muscu
19	1178.5	48.8	813	14 CD519131	CD519131 AGENCOURT
20	1166	48.2	708	14 CD366462	CD366462 UI-H-FT1-
21	1164.5	48.2	922	12 BG761741	BG761741 602717936
22	1154	47.7	945	12 BE909780	BE909780 603070868
23	1144.5	47.4	889	10 BE906771	BE906771 601498329
24	1130.5	46.8	1081	13 BQ072892	BQ072892 AGENCOURT
25	1112	46.0	1020	13 BX439466	BX439466 BX439466
26	1105	45.7	801	12 BI754143	BI754143 603025609
27	1092	45.2	701	12 BI771109	BI771109 603059866
28	1092	45.2	783	12 BI754468	BI754468 603022607
29	1086	44.9	650	10 BG432644	BG432644 602466315
30	1068	44.2	786	10 BG430966	BG430966 602500267
31	1063.5	44.0	944	12 BI819273	BI819273 603037713
32	1059	43.8	879	13 BQ218838	BQ218838 AGENCOURT
33	1053	43.6	653	10 BE218907	BE218907 hv46D06.x
34	1052.5	43.5	1073	13 BX360506	BX360506 BX360506
35	1043.5	43.2	822	12 BI911772	BI911772 603065377
36	1036.5	42.9	805	10 BG563740	BG563740 602584524
37	1031	42.7	677	10 BE670671	BE670671 7638606.x
38	1014	42.0	635	14 CB448111	CB448111 702171 MA
39	1000	41.4	736	12 BI769575	BI769575 603054947
40	996	41.2	676	9 AA796802	AA796802 VP1D01.r
41	995	41.2	751	14 CD558931	CD558931 AGENCOURT
42	990	41.0	637	13 BQ564473	BQ564473 5117610.Y
43	983.5	40.7	849	10 BG166321	BG166321 602345529
44	980	40.5	806	12 BI764052	BI764052 603043258
45	979	40.5	727	12 BG765433	BG765433 602738911

## ALIGNMENTS

RESULT 1  
LOCUS AK032972  
DEFINITION Mus musculus 12 days embryo male wolffian duct includes surrounding region cDNA, RIKEN full-length enriched library, clone:672048C06  
product:plasma glutamate carboxypeptidase, full insert sequence.  
ACCESSION AK032972.1 GI:26328732  
VERSION AK032972.1  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
99279253  
10349636

1  
Carninci P., Hayashizaki Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
20499374  
11042159

2  
Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K.,  
Itoh M., Kono H., Okazaki Y., Muramatsu M. and Hayashizaki Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
2530913  
11076861

3  
Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
Kono H., Akiyama J., Nishi K., Katsuna T., Tashiro H., Itoh M.,  
Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
Fujikawa S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,  
Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
Okazaki Y., Muramatsu M., Inoue Y., Kita A. and Hayashizaki Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
21085660  
11217851

4  
Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
Aizawa K., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,  
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
Kadota K., Matsuda H., Ashburner M., Barlow S., Casavant T.,  
Pletschmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
Kuehl P., Lewis S., Matsuo Y., Nakaido I., Pesole G.,  
Queckenbush J., Schriml L.M., Staudil F., Suzuki R., Tomita M.,  
Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H.,  
Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N.,  
Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C.,  
Fletcher C., Fujita M., Gariboldi M., Gastlrich S., Hill D.,  
Hochmann M., Hume D.A., Kamita M., Lee N.H., Lyons P.,  
Machlomi L., Washima J., Mazzairelli J., Momberts P., Nordone P.,  
Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H.,  
Sato K., Schonbach C., Seva T., Shibata Y., Storch K.F., Suzuki H.,  
Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmberg L.,  
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kontuski S.  
and Hayashizaki Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
21085660  
11217851

5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
21085660  
11217851

6  
Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
Kato H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M.,  
Koye S., Kurihara C., Matsuyama T., Miyazaki A., Murata M.,  
Nakamura M., Nishi K., Nomura K., Nunazaki R., Ohno M., Ohsato N.,  
Okazaki Y., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N.,  
Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T.,  
Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S.,  
Takeda Y., Tanaka T., Tomaru A., Toyo T., Yasunishi A.,  
Muramatsu M. and Hayashizaki Y.  
Direct Submission  
Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome

COMMENT  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in Riken  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/.

FEATURES  
SOURCE  
Location/Qualifiers  
1..1739  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:6720483C06"  
/db\_xref="taxon:10090"  
/clone="6720483C06"  
/sex="male"  
/tissue\_type="olfactory duct includes surrounding region"  
/clone\_id="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="12 days embryo"  
81..1493  
/note="unnamed protein product; plasma glutamate  
carboxypeptidase (MCD) [MGI:1889205, GB|NM\_018735, evidence:  
BLASTN, 99%, match=943]  
putative"  
/protein\_id="BAC28105.1"  
/db\_xref="GI:26328733"  
/db\_xref="FANTOM DB:6720483C06"  
/transcript="MUSFLPIVHLALGSGKAVPKNQVRSQRTFREKEIANEYV  
AKATINAVGKYONSRYERLGLVDVGRSLGSKULERAIDIMONLQDGLBNH  
LEQVRIPIHMGESAVMLPRIRKAIILGSSIGTPPGITAEVLVVASFPELQR  
ASEBGRKIIYVNOPTGYEVQYQAVAAKVGAVASLIQSVAFSISPTGICR  
KYDGVGKIPACTIVEDAMSRGKRIYHLEMAKTYPDTSFNTVAITSS  
MYPEVVLVSGHIDSWDGGALDGGGATISWALSTVADIGLRPRTLRVLMTE  
EKGICGASQYELHKKANISKYSLYMEADSTFPTGICQFGSDKARIMKENVLLP  
LNVATKVSNEGDIVINIQAGVGSALRDVLKYFFHSHSDTMTMPKONNVAA  
AAWAAVVAADDEMILPRS"  
1722..1727  
/note="putative"  
1739  
/note="putative"  
polya\_signal  
polya\_site  
BASE COUNT 482 a 369 c 423 g 465 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,47e-235 Length: 1739  
Score: 2135.00 Matches: 414  
Percent Similarity: 93.64% Conservative: 28  
Best Local Similarity: 87.71% Mismatches: 28  
Query Match: 88.25% Indels: 2  
DB: 11 Gaps: 1  
US-09-745-763-36 (1-472) x AK032972 (1-1739)  
QY 1 MetLysPheLeuIlePheAlaPheGlyGlyValHisLeuLeuSerLeuGlySerGly 20  
DB 81 ATGAGGCCCTTTCTTCCTCGTTTCAT-----GTTACCTTTAGCCTTGCGCTCGT 134  
QY 21 LysAlaIleCysLysAlaGlyIleSerLysAlaGlyThrPheGluGluIleLeuGluGlu 40  
DB 135 AAACCTGATTCAGAAATGGTGTTCCACGCAACATTCGAGAAATTAAGAAATA 194  
QY 41 Alasercysglaspvalalalyalailelelsenuulavalytyrglylvaslagln 60  
DB 195 GCCACATGAGAGTGTGCTAAAGCAATATCACTTCGTTTATGGTAATATACAG 254  
QY 61 AsnArgSerTyrGluGluAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80

[illegible][illegible]

REFERENCE  
AUTHORS

TITLE

JOURNAL

REFERENCE  
AUTHORS

5  
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 1805)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arikawa, T., Bono, H., Caminici, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komori, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Nunahara, R., Ohno, M., Okasaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Yamatsuta, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-Apr-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers  
1. 1805  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM:DB:1190003P12"  
/db\_xref="MGI:1909852"  
/db\_xref="Taxon:10090"  
/clone="1190003P12"  
/tissue\_type="whole body"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="18-day embryo"  
147. 1559  
/note="unnamed protein product; plasma glutamate carboxypeptidase (MGP[MGI:1889205, GB|NM\_018755, evidence BLASTN, 99%, match=943)  
putative"  
/codon\_start=1  
/protein\_id="BAC35891.1"  
/db\_xref="GI:26344481"  
/db\_xref="MGI:1923585"  
/translation="MRSLFLPLFIVHLALGSKAVKNGVSORPREIKERIANEYEDAKATINLAVGVKYNRSYERELGVDTVGPLRSGKNLEKIMQIMYQLODGLENVLALEQRIPIHMERGESAAVLEPRIHKALILGAGSGCTPGGITAELVVAFAFDELQRAASEAGKTIIVNQPTVEGKYVQVQAVEAKGVAASLIQSVASFIISPHGICRYQDDVPKIPACTIVEDAEMWSPKAGNKIVHLENGAKTYPTDSFNNVAETGMYPERVAVSGGLSDMDGQGLDGGCAFTSMELSLVKDGLRPRKTLVLVYTAABECCGGLASQVGYELHKANSKYSKLWAEASGLFISGLQFTGSDKRAIMKEVMNLLQELNIVTFPSNGEGTDINFWIQAGVPSASLRDLYKFFPHSHGDMTMDKQNNVAAAVMAVVAIVVADMDMELPRS"  
1788. 1793  
/note="putative"  
1805  
/note="putative"  
BASE COUNT 486 a 397 c 445 g 476 t 1 others  
ORIGIN  
Alignment Scores:

Pred. No. :	1,55e-235	Length:	1805
Score:	2133.00	Matches:	414
Percent Similarity:	93.64%	Conservative:	28
Best Local Similarity:	87.71%	Mismatches:	28
Query Match:	88.25%	Indels:	2
DB:	11	Gaps:	1

US-09-745-763-36 (1-472) x AK075686 (1-1805)

QY	1	MethylserLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGly	20
DB	147	ATGAGGCTCCCTTTCTTCTCTGTCATT-----GTTCACTTTTAGCTTGCGGCTCTGCT	200
QY	21	LysAlaIleCybLybAserGlyTLeSerLybArgTrPheGluGluIleLybGluGluIle	40
DB	201	AAAGCTGATTCCAGAAATGGTGTTCACGGAACATTTCAGAAATTAAGAAATA	260
QY	41	AlaserCybLybAspValAlaLysAlaIleLeuLeuAlaValTyrGlyLybAlaGln	60
DB	261	GCCAACTATGTAAGATGTGCTTAAGCCATTATCAACCTTGCTTTATGTGTAATACAG	320
QY	61	AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly	80
DB	321	AACGGCTCTATGAGCGCTTGGGACTTCTAGTTGATCTGTTGGACCCAGACTGAGTGC	380
QY	81	SerLybAsnLeuGluLybAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu	100
DB	381	TCTAAGAACCTTAGAGAAAGCTATTCAATCTATGTAACCAAACTGCGACCAAGATGGCTA	440
QY	101	GluLybValHisLeuGluProValArgTLeProHisTrpGluArgGlyGlnIleSerAla	120
DB	441	GAAATGTTCACTCGAGACAGCTCAGATCAATCCCACTGGAGAGGGAGAAAGATCTCA	500
QY	121	ValMetLeuGlnProArgIleHisLybIleAlaIleLeuGlyLeuGlySerSerIleGly	140
DB	501	GTGATGTCTTGAAGCTTCGAATTCAACAAGTGGCTATTCTAGGCTCTTGGCAGACGATTTGG	560
QY	141	ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln	160
DB	561	ACTCCCCCAGAGGAGCATCAGACAGAAAGTGGTGGTGGAGGCTCTTTGATGAACCTTCA	620
QY	161	ArgArgAlaSerGluAlaArgGlyLybSerIleValValTyrAsnGlnProTyrIleAsnTyr	180
DB	621	AGAAGAGCATCAGAAAGCAAGAGGAGGATCTATTGTTAATACAGACCTTACACTGGCAT	680
QY	181	SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaAlaLybValAlaLeu	200
DB	681	GAGAAAGCTGTGCAGTACCGGGGTGCAGGAGACTGTGGAAGCTGCCAAGTGGAGACTGTG	740
QY	201	AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln	220
DB	741	GCATTCCTCATTCACATCACTAGTCTTCTTCTTCATCTACAGTCTCTCACACAGGATTCAG	800
QY	221	GluTyrGlnAspGlyValProLybIleProThrAlaCybIleThrValGluAspAlaGlu	240
DB	801	AAATATCAAGATGGGTGCCCAAGATTCCAAAGCTGTATCAACAGTAAAGATGACADA	860
QY	241	MetMetSerArgMetAlaSerHisGlyIleLybSerIleValIleGlnLeuLybMetGlyAla	260
DB	861	ATGATGTCTCCAAATGGCTTCTCGTGGGAAACAAATTGTCAATTCATCGGAATGGGACA	920
QY	261	LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLybTyr	280
DB	921	AAGACTATCCAGATCTGATTTCTTCAATCTGTTCAGAGATCACTGGAGACATGAT	980
QY	281	ProGlnGlnValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla	300
DB	981	CCAGAGGAAGTGTCTGTGCAGTGGACATTGGACAGCTGGAGTGTGGGCAAGGTGCA	1040
QY	301	MetAspAspGlyGlyGlyAlaPheIleSerTrpGluAlaLeuSerLeuIleLybAspLeu	320
DB	1041	CTGATATATGTGTGTGAGCTTATATATCAAGGAAAGCACTCACTTGTAAAGATCTT	1100



QY	321	GLYLEUENRGP	PROLYSARG	THYLEMARG	LEUVALLEUT	TPTRAL	AGLUGLNG	GLYGLY	340														
Db	1101	GGCGTCGCT	CAAGAGGAC	CTTCGCCCT	GGTGGCT	CTTGAC	TCGAAGA	CAAGAGAGA	1160														
QY	341	VALGLYAL	APHEGIN	TYR	TYRGLN	LEUHI	ISLYS	VALAENI	LESER	NTY	SER	LEU	VAL	360									
Db	1161	ATYGGTGCT	CCCAAGTAT	TATAGAG	CTCAT	TAAGGCA	AAATAT	TTTCCAA	GTACAG	CTTGATG	1220												
QY	361	METGLUSE	APPA	AGLY	THR	PHLEUP	PRO	THR	GLYLEU	GIN	PHI	THR	GLY	SER	GLN	LYS	380						
Db	1221	ATGGAAGCT	GACTCA	GGAACCT	CTTTAC	CCACTG	ACGTC	GACGTT	CAC	GTTCAG	SGAGGAC	CAAG	1280										
QY	381	ALAARG	ALALEM	ETGLU	VAL	VALMET	SER	LEU	LEU	GIN	PRO	LEU	AENI	LE	THR	GLN	VAL	400					
Db	1281	GCCAGGGCT	ATCATGA	AGAGAA	GAGTAT	GAATCT	TTCG	CAACCC	CTCAAT	GTACCA	AGCT	1340											
QY	401	LEUSER	HIS	GLY	GLU	GLY	THR	ASP	ILE	ASN	PHET	IP	LE	GIN	AL	AGLY	VAL	PRO	GLY	VAL	420		
Db	1341	TTTGTATAT	GAGAGA	GAAGCA	AGCAT	TATTA	CACTTC	GGAC	CAACGCT	GAGGTAC	CTGGAGCC	1400											
QY	421	SER	LEU	LEU	ASP	AP	LEU	TYR	LYS	TYR	PH	EP	HE	PH	EN	HIS	SER	HIS	AGLY	ASP	THR	MET	440
Db	1401	AGTCTGCT	GATGACT	GTTCAC	AGAT	TTCTTT	TTCAT	CACTT	CCACG	AGAC	CAACCATG	1460											
QY	441	THR	VAL	MET	ASP	PRO	LYS	GLN	MET	ASN	VAL	ALA	ALA	VAL	ITP	ALA	VAL	VAL	SER	TYR	460		
Db	1461	ACTGTCAT	GATGAT	CCAAAC	AGATGA	TATTTG	CTGCTG	CTGAT	GTATGG	GGCTG	TGTG	CGCTTAT	1520										
QY	461	VAL	VAL	ALA	ASP	MET	GLU	GLN	MET	LEU	PRO	ARG	SER	472									
Db	1521	GTTGTG	CAGACAT	GATGA	ATAATG	TGCTG	CCCAAG	TTC	1556														

```

REMARK
COMMENT
FEATURES
SOURCE
1. 1868
/organism="Mus musculus"
/mol_type="mRNA"

```

DB	QY	Alignment Scores:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:
		1.79e-234	2124.00	94.79%	88.94%	87.88%	
		Length: 1868	Matches: 410	Conservative: 27	Mismatches: 24	Indels: 0	
		Gaps: 0					
US-09-745-763-36 (1-472) x BC010977 (1-1868)							
QY	12	ValHisLeuLeuSerLeuCysSerGlyLysAlaIleCysLysAsnGlyIleSerIleArg	31				
Db	228	GTTCACTCTTTAGCCTTGCGGCTCTGAGTAAAGCTTATTCAGAAATGGTGTCTTCACGCGA	287				
QY	32	ThrPheGluGluIleLysGluGluIleAlaSerCysGlyAspValAlaLysAlaIleIle	51				
Db	288	ACATTTCGAGAAATTAAGAAGAAATATGCCACATATGATGATGCTTGAAGCAATTATC	347				
QY	52	AsnLeuAlaValIlyrGlyLysAlaGlnAsnArgSerTyrGluArgLeuAlaLeuLeuVal	71				
Db	348	AACCTGTGCTTTTGTGTAATATACAGAAACCGGCTCTATGAGCCCTTGGGACCTCTAGTT	407				
QY	72	AspThrValGlyProArgLeuSerGlySerIleAsnLeuGluLysAlaIleGlnIleMet	91				
Db	408	GATACGTGTGAGCCAGCTGAGTGCGCTTAAAGACCTAGAGAAAGCTATCCAATCATG	467				
QY	92	TyrGlnAsnLeuGlnGlnAspGlyLeuGluLysValHisLeuGluProValArgIlePro	111				
Db	468	TACCAAAACCTCGACGACGAAGATGGGCTAGAAATTTTCACTCGGACGAGCTCAGATACC	527				
QY	112	HisTrpGluArgGlyGluGluSerAlaValMetLeuGluProArgIleHisIleLeuAla	131				
Db	528	CACGCGAGAGGGAGAAAGATCTGCAGTGTGCTTGAGCTCGCAATTCACAAAGATGGCT	587				
QY	132	IleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIleThrAlaGluValLeu	151				
Db	588	ATTCAGAGCTTGGACAGCAGCATTTGGGACTCCCCAGAGGCATCACAGCGAAGTGGTG	647				
QY	152	ValValThrSerPheAspGluLeuGlnArgArgAlaSerGluAlaArgGlyLysIleVal	171				
Db	648	GTCGTAGCCTCTTTTGATGAATCTCAAGAAAGACATCAGAAAGCAAGAGGAAGATCATT	707				
QY	172	ValTyrAsnGlnProTyrIleAsnTyrSerArgThrValGlnTyrArgThrGlnGlyAla	191				
Db	708	GTTATATACACGACCTTACACTGCGCTATGAGAAAGCTGTCAGTACCGGGTCCAGGAGACT	767				
QY	192	ValGluAlaAlaLysValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPheSer	211				
Db	768	GTCGAGCTGCAAGTGGAGGCTGTGGCACTCCCTCATCCAAATAGTAGCTTCCTTTC	827				
QY	212	IleTyrSerProHisThrGlyIleGlnGluTyrGlnAspGlyValProLysIleProThr	231				
Db	828	ATCTACAGCTCTTCACACAGGATTTTACAGAAATATCAAGATGGTGTGCGCCAGATTCACACA	887				
QY	232	AlaCysIleThrValGluAspAlaGluMetLysSerArgMetAlaSerHisGlyIleLys	251				
Db	888	GCCGTATACACAGTAAAGATGACGAATATGATGCTCGAATGGCTTCGCGGGAACAA	947				
QY	252	IleValIleGlnLeuLysMetGlyAlaLysThrTyrProAspThrAspSerPheAsnThr	271				
Db	948	ATTGTCATTCATCTGCGAAATGGAGCAAAAGCCATACAGATGACTGATTCCTTCAATACT	1007				

272 ValAlaGluIleThrGlySerLeuTyrProGluGlnValIleuValSerGlyHisLeu 291  
 1008 GTTGAGAGATCAGTCCGAGAGCATGTATCCAGAGAACTTCTGCTCAGTGCAGATTTG 1067  
 292 AspSerTyrAspValGlyGlnGlyValIleuAspAspGlyGlyGlyValAlaPheIleSerTyr 311  
 1068 GACAGCTGGAGATGTTGGGAGAGGTCCTCGATGATGGTGGTGGAGCTTATCATCAGG 1127  
 312 GluAlaLeuSerLeuIleLeuAspLeuGlyLeuArgProGlyArgThrIleuArgLeuVal 331  
 1128 GAAGACCTCTCACTTGTAAATATCTTGGGCTGCTCCAAAGAGAACTCTGGCTGGTGG 1187  
 332 LeuTyrThrAlaGluGlnGlyGlnGlyValGlyAlaPheGlnTyrTyrGlnLeuHisLeu 351  
 1188 CTCGTGACTGCGAAGAAACAGAGAGATGTGTCTCTCCAGTATATGAGTACATTAAG 1247  
 352 ValAlaIleSerAspTyrSerLeuValMetGluSerAspAlaGlyTyrPheLeuProThr 371  
 1248 GCAATATTTTCCAGTACAGCTTGGTGAATGAGAACTGACAGAACCTTCTTACCACCT 1307  
 372 GlyLeuGlnPheThrGlySerGlyValAlaArgAlaIleMetGluGlnValMetSerLeu 391  
 1308 GGAGTCGAGTTCACGTGAGAGTGAAGAGCCAGGCTATCATGAGAGAACTCATGATCTT 1367  
 392 LeuGlnProLeuAsnIleThrGlnValIleuSerHisGlyGlyGlyTyrAspIleAsnPhe 411  
 1368 CTGCAACCCCTCAATGTCACCAAGGTCTTTAGTATGAGAAAGAACTGATTTTAACTTC 1427  
 412 TrrIleGlnAlaGlyValProGlyValIleSerLeuLeuAspAspLeuTyrIleTyrPhePhe 431  
 1428 TGGATCCAGTCAAGTGAAGTACCTGAGCCAGTCTGCCGATGATGATGAGAACTTCTT 1487  
 432 PheHisLeuSerHisGlyAspThrMetThrValIleMetAspProGlyGlnMetAsnValAla 451  
 1488 TTCCATTCATTCCTCAGAGAGACCAATGATGATGATGAGAAAGAACTGATTTTAACTTC 1547  
 452 AlaAlaValTrrPalaValIleValSerTyrValValAlaAspMetGlyGlnMetLeuProArg 471  
 1548 GCTGCTGTGGGCTGTTGTCTGCTTATGTTGTTGTCAGACATGAGTGAATGTTGCCACAG 1607  
 472 Ser 472  
 1608 TCC 1610  
 RESULT 4  
 BC012019 1024 bp mRNA linear HTC 04-MAR-2003  
 LOCUS Homo sapiens, similar to plasma glutamate carboxypeptidase, clone  
 DEFINITION IMAGE:4455631, mRNA.  
 ACCESSION BC012019.1 GI:15080558  
 VERSION 1  
 KEYWORDS Homo sapiens (human)  
 SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 1024)  
 Strausberg, R.  
 DIRECT SUBMISSION  
 TITLE Submitted (30-JUN-2001) National Institutes of Health, Mammalian  
 JOURNAL Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 CONTACT: MGC help desk  
 Email: [cgapps-r@mail.nih.gov](mailto:cgapps-r@mail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louieged, H.,  
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Naranvali,  
 A.N., Gibbs, R.A.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>  
 Series: IRAX Plate: 28 Row: b Column: 11  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 5174626  
 This clone has the following problem: retained intron.  
 Location/Qualifiers  
 1..1024  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4455631"  
 /tissue\_type="Kidney, hypernephroma"  
 /clone\_lib="NIH MGC\_89"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-Sport6"  
 BASE COUNT 317 a 210 c 256 g 241 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,23e-178 Length: 1024  
 Score: 1633.00 Matches: 323  
 Percent Similarity: 99.09% Conservative: 2  
 Best Local Similarity: 98.48% Mismatches: 1  
 Query Match: 67.56% Indels: 2  
 DB: 11 Gaps: 1  
 US-09-745-763-36 (1-472) x BC012019 (1-1024)  
 1 MetIysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCysSerGly 20  
 36 ATGAATTCCTTATCTTCGCACTTTTGGTGGTTCACCTTTATCCCTGCTGCTGGG 95  
 21 LysAlaIleCysLysAsnGlyIleSerTyrArgThrPheGluGluIleLeu 40  
 96 AAAGCTATATGCAAGATGCGATCTTAAGAGGACTTTTGAAGAAATTAAGAAATA 155  
 41 AlaserCysGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60  
 156 GCCAGCTTGGAGATGTTGCTTAAGCAATCACTACCTGCTTTATGCTTAAGCCAG 215  
 61 AsnArgSerTyrGluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80  
 216 AACAGATCTTATGAGCATGTCGCACTTCTGTTGATCTGTTGAGACCACTGAGTGGC 275  
 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAsnGlyLeu 100  
 276 TCCAGAGACCTTAAGAAAAAGCCATCAATTAATTAAGCAAACTGCGAGCAAGATGGCTG 335  
 101 GluLysValHisLeuGluProValArgIleProHisTrrGluArgGlyGlnGlySerAla 120  
 336 GAGAAAGTTCACTCGAGCCAGTGAATACCACTCGAGAGAGGAGAGAAATCACT 395  
 121 ValMetLeuGluProArgIleHisLysAlaIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
 396 GTGATGCTGAGCAAGAAATCATTAAGTATACCACTCGGCTTGGCAGCAGCATTTGGG 455  
 141 ThrProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160  
 456 ACTCTCCAGAGGACATTACAGCAAGAGTTCTGGTGGAGCTCTTTCAGTGAACCTGAG 515  
 161 ArgArgAlaSerGluAlaArgGlyLysAlaIleValValTyrAsnGlnProGlyTrrAsnTyr 180  
 516 AGAAGGGCTTGAAGAGCAAGAGGAGATTTGTGTTTAAACCAACTTATACATCACTAC 575  
 181 SerArgThrValGlnTyrArgThrGlnGlyValAlaGluAlaAlaLysValGlyAlaLeu 200

Db	576	TCAAAGACGGTGCATATACCGAAGCGAGGGGGCGGTGAAAGCGTCCAAAGCTTTGGGCTTTG	635
Qy	201	AlaSerLeuIleIeaGSeValAlaIaSerPheSerIleIyrSerProHisIleThrGlyIleGln	220
Db	636	GCATCTCTCATTCGATCCGTGGCCCTCTTCTCATCTACGATCTCAACAGAGTATTCAG	695
Qy	221	GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGlnAspAlaGlu	240
Db	696	GAATACCCGAGTGGCGTCCCAAGATTCCAACAGCGCTGATTAACGTTGGAAGATCAGAA	755
Qy	241	MetMetSerIaGMeIaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyVala	260
Db	756	ATGATGTTCMAAAATGGCTTCTTCATGGGATTCAAAATGTTCATTCAGCTAAAGATGGGGCA	815
Qy	261	LysThrIyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIleIyr	280
Db	816	AAGACCTAACCCAGAACTGATCTCTTCAACTGTAGAGAGATCACTGGGAGCAAAATAT	875
Qy	281	ProGluGlnValIleValLeuValSerGlyHisLeuAspSerTPAspValGlyGlnGlyVala	300
Db	876	CCAGAAACGGTTGATCTGTCACTGGAGACATCTGGACACCTGGAGATGTTGGCGAGGGTCC	935
Qy	301	MetAspAspGlyGlyGlyValaPheIleSerTPGluAlaLeuSerLeuIleLysAspLeu	320
Db	936	ATGGATGTATGGGGGTGGAGCCCTTATATATCATGGGAAGCATCTCACTTATTAAGATCTT	995
Qy	321	Gly-----LeuArgProLysArg 326	
Db	996	GGTAATATTTTAAAAAATAAAAAA 1019	

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BX439467	1050 bp mRNA linear	BX439467	BX439467	5-PRIME, mRNA sequence.	EST.	Homo sapiens (human)	1 (bases 1 to 1050)	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	Full-length cDNA libraries and normalization	unpublished	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: sege@genoscope.cns.fr Web : www.genoscope.cns.fr Library was constructed by life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4663.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0DE010DE30P1&cluster=4663.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DE010DE30P1.

FEATURES	SOURCE
Location/Qualifiers	1..1050
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="CSODE010YJ06"	
/tissue_type="PLACENTA"	
/clone_id="Homo sapiens PLACENTA"	
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."	
BASE COUNT	307 a 233 c 271 g 239 t
ORIGIN	

Alignment Scores:	
Pred. No.:	4.1e-167
Score:	1539.00
Percent Similarity:	100.00%
Best local Similarity:	100.00%
Query Match:	63.67%
DB:	13
US-09-745-763-36 (1-472) x BX439467 (1-1050)	
	Length: 1050
	Matches: 303
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0

Oy	1	MetIysPheLeuIleIlePheAlaPhePheGlyValIleHisLeuLeuSerLeuCysSerGly	20
Dd	142	ATGAATTCCTTAATCTTCGACATTTTTCGGTGTGTTCACCTTTTATCCCTGTGCTCTGG	201
Oy	21	LYsAlaIleCysGlysaenglyIleSerIysaArgThrPheGluGluIleGlyGluIle	40
Dd	202	AAACCTATATSCAAGAATGGCATCTCTTAAGAGGACCTTTTGAAGAAATAAAGAAATA	261
Oy	41	AlaSerCysGlyaaPheValAlaIleIleLeuLeuAlaValIleGlyIleValaGln	60
Dd	262	GCCAGCTGTGAGATGGTGTCAAGCAATCAACTAGCTGTTATAGTAAACCCAG	321
Oy	61	AsnArgSerIleYrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly	80
Dd	322	AACGATCTCTATGAGCCATTTGGCATCTTGTGTATCTGTTGAGACCCAGACTAGTGGC	381
Oy	81	SerIysaLeuGluIleGluValaIleGlnIleMetIyrGlnaLeuGlnIleAspGlyLeu	100
Dd	382	TCCAAGAACTAGAAAAAGCCATCAATATTATGACCAAACTGACACAGATGGGCTG	441
Oy	101	GluIysValHisLeuGluProValArgIleProHisIleTyrGluArgGlyValGluSerIle	120
Dd	442	GAGAAAGTCACTCGAGGCCAGTAGAATATCCCACTGGGAGAGGGGAGAAAGAACACT	501
Oy	121	ValMetLeuGluProArgIleHisIleIysIleAlaIleLeuGlyIleGlySerSerIleGly	140
Dd	502	GTGATGTCTGGAGCCAGAAATTCATTAAGATACCATCTCGGGTCTTGGAGCAGCATTTGG	561
Oy	141	ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln	160
Dd	562	ACTCTCTCAGAAAGCATTAACAGAGAAAGTTCTGGGTGAGACTCTTTTGATGAACTGGAG	621
Oy	161	ArgArgAlaSerGluAlaArgGlyIleValIleValIleTyrAsnGlnProTyrIleAsnIyr	180
Dd	622	AGAAAGGGCTCAGAAAGCAAGAGGAGATGTGTGTTATTAACCAACTTACATCAACTAC	681
Oy	181	SerArgThrValGlnIyrArgThrGlnGlyAlaValGluAlaAlaIysValGlyAlaLeu	200
Dd	682	TCAAGAGACGGTGCATTAACCGAAGCAGGGGGCGGTGGAAGCTGCCAAGTTGGGGCTTGG	741
Oy	201	AlaSerLeuIleArgSerValAlaIleSerPheSerIleIyrSerProHisIleThrGlyIleGln	220
Dd	742	GCATCTCTCATTCATTCGTTGGCTCTCTTCCATCTACAGTCCCTCACACAGGTATTTCAG	801
Oy	221	GluIyrGlnaAspGlyValProIysIleProThrAlaCysIleThrValGlnaAspAlaGln	240
Dd	802	GAATTAACCGATGGCGTGGCCCAAGATTCCAACGCTGATTAACGGTGAAGATCAGAA	861
Oy	241	MetMetSerArgMetAlaSerHisGlyIleIysIleValIleGlnLeuIysMetGlyAla	260
Dd	862	ATGATGTCAAGATAGGCTTCTCATGCGATCAAAATTTGCATTCAGCTAAAGATGGGGCA	921
Oy	261	LysThrIyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIyrIyr	280
Dd	922	AAGACCTTAACCAAGTACTGATTTCTTCAACTGTAGAGAGATCACTGGGAGCAAAATAT	981
Oy	281	ProGluGlnValIleValLeuValSerGlyHisIleuAspSerTyrAspValGlyGlnIlyAla	300
Dd	982	CCAGAACAGGTGTGACTGTGTCAGTGAATCTTGACACACTGGGATGTTGGGAGGGGTGCC	1041
Oy	301	MetAspAsp	303

```

Db      1042 ATGATGAT 1050

RESULT 6
LOCUS   BX355940
DEFINITION BX355940 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CS0D1006Y115 5-PRIME, mRNA sequence.
ACCESSION BX355940
VERSION  BX355940.1 GI:30384019
KEYWORDS EST.
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1038)
AUTHORS  Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE    Full-length cDNA libraries and normalization.
JOURNAL  Unpublished
COMMENT  Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4663.f for more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1006A80P1&cluster=4663.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0D1006A80P1.
LOCATION/Qualifiers
1..1038
/mol_type="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1006Y115"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT  304 a ~237 c 262 g 233 t
ORIGIN

Alignment Scores:
Pred. No.: 1,28e-154 Length: 1038
Score: 1431.00 Matches: 286
Percent Similarity: 99.65% Conservative: 0
Best Local Similarity: 99.65% Mismatches: 1
Query Match: 59.21% Indels: 1
DB: 13 Gaps: 0

US-09-745-763-36 (1-472) x BX355940 (1-1038)

Oy      1 MetLysPheLeuI1PheAlaPhePheGlyValHisLeuLeuSerLeuCySerGly 20
Db      186 ATGAATTCCTATCTTCGCACTTTTCGGTGTGTTACCTTTATCCCTGCTCTGG 239
Oy      21 LysAlaIleCyValAsnGlyIleSerIysArgThrPheGluGluIleGluGluIle 40
Db      240 AAAGCTATATGCAAGATGCGATCTCTAAGAGCATTTTGAAGAATATAAGAAATA 299
Oy      41 AlAserCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValItyGlyValAgin 60
Db      300 GCCACCTGTGGAGATGTCTAAACCAATCAACCACTAGCTTTATGTTAAACCCAG 359
Oy      61 AsnArgSerTyrGlyArgLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80
Db      360 AACCAATCTTATGACGATGCGACTTCTGTTGATACCTGTTGGACCCAGACTGAGTGC 419
Oy      81 SerLysAsnLeuGluLysAlaIleGluIleMetTyrGlnAsnLeuGlnAspGlyLeu 100

```

```

Db      420 TCCAAGAACTAGAAAAAGCCATCCAAATTATGTACCAAAACTGTGACAGATGGGCTG 479
Oy      101 GluLysValHisLeuGluProValArgIleProHisIleProGlyValGluSerAla 120
Db      480 GAGAAAGTTCACTCGAGGCCAGTGAATATCCCACTGGGAGAGGGAGAAATCAAGCT 539
Oy      121 ValMetLeuGluProValGlyIleHisLeuIleAlaIleLeuGlyLeuGlySerSerIleGly 140
Db      540 GTGATGCTGAGACCAAGATTCATTAAGATAGCCATCTCGGCTCTTGGACAGCATGGG 599
Oy      141 ThrProGluGlyIleThrAlaGluValLeuValIleThrSerPheAspGluLeuGln 160
Db      600 ACTCTCCAGAAAGCATTACAGCAAGATTCTGCTGTGAGACCTCTTGATGAACCTCAG 659
Oy      161 ArgArgAlaSerGluAlaArgGlyValIleValValIleThrAsnGlnProTyrIleAsnTyr 180
Db      660 AGAAGGCTCTAGAAAGCAAGAGGAAAGTGTGTATTAACCAACTTACATCACTAC 719
Oy      181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaValAlaValAlaLeu 200
Db      720 TCAGAGACGTGCAATACCAAGACGAGGGCGGTGGAGCTGCAGAGTTGGGGCTTTG 779
Oy      201 AlAserLeuIleArgSerValAlaSerPheSerIleTyrSerProHisIleGln 220
Db      780 GCATCTCATTCGATCCGTGGCTCTCTCCATCTACAGTCCCTCACACAGGATTCAG 839
Oy      221 GluTyrGlnAspGlyValProValIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db      840 GAATACCAAGATGAGCGGCCCAAGATTCCAACACCTGATTAAGGAGGAAAGTCAGAA 899
Oy      241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMet-GlyAla 260
Db      900 ATGATGTCAGAAATAGGCTTTCATGGGATCAAAATTCATTCAGTCAAGATCGGGGCG 959
Oy      260 alyGThrTyrProAspThrPheSerPheAsnThrValAlaGluIleThrGlySerIty 280
Db      960 AAACACCTACCAAGATCTGATTCCTTCAACACTGACAGATCACTGGAGCAATA 1019
Oy      280 rProGluGlnValValLeu 286
Db      1020 TCCAAGACGATGTACTG 1038

RESULT 7
LOCUS   B0878966
DEFINITION AGENCOURT_8183107 Lupski_dorsal_root_ganglion Homo sapiens CDNA clone IMAGE:6184129 5', mRNA sequence.
ACCESSION B0878966
VERSION  B0878966.1 GI:22270974
KEYWORDS EST.
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 909)
AUTHORS  NIH-MGC http://mgs.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished
JOURNAL  Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LRAM3572 row: n column: 02
High quality sequence stop: 696.
LOCATION/Qualifiers
1..909
/mol_type="Homo sapiens"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:6184129"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_id="lupski_dorsal_root_ganglion"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
directionally cloned using the following adaptors:
5'-TCGACCCACGCGCTCG-3' and
5'-GACTAGTTCATGATCGCGAGCGCGCCCTT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
BASE COUNT      273 a      226 g      207 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:      1,48e-145      Length:      909
Score:          1352.00      Matches:      276
Percent Similarity: 97.89%      Conservative: 2
Best Local Similarity: 97.18%      Mismatches: 3
Query Match:    55.94%      Indels:      3
DB:            13      Gaps:      0
US-09-745-763-36 (1-472) x B0878966 (1-909)
QY      1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGly 20
Db      64 ATGAATTCCTTATCTTCGATTTTTCGGTGTGTCACCTTTATCCCTGCTGCGG 123
QY      21 LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleGluGluIle 40
Db      124 AAACCTATGCAAGAAATGCGATCTCTTAAGAGGACTTTTGAAGAAATTAAGAAATA 183
QY      41 AlasrCySglYarpValAlaLysAlaIleLeuLeuAlaValIleTyrGlyLysAlaGln 60
Db      184 GCCAGCTGTGAGAGATGTGCTTAAGCAATCAACCACTGATTATGTAAGGCCAG 243
QY      61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
Db      244 AACGATCTATGAGCGCATTCGCACTTGTGTTATCTGTGACCCAGACTGATGCG 303
QY      81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnInAspGlyLeu 100
Db      304 TCCAAGAACTTAGAAGAAAGCCATCCAAATTAATGTAACAAACCTGACAGATGCGCTA 363
QY      101 GluLysValHisLeuGluProValArgIleProHisTTPGluArgGlyGluGluSerAla 120
Db      364 GAGAAAGTTCACCTGAGCCAGTGAATACCCCACTGGAGAGGAGGAAAGAAATCAAGCT 423
QY      121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
Db      424 GTGATGCTGGAGCCAGAAATTCATAGATAGCCATCTCGGCTTTGGACAGACATTTG 483
QY      141 ThrProGluGlyIleThrAlaGluValLeuValIleThrSerPheAspGluLeuGln 160
Db      484 ACTCCTCCAGAAAGGATTACAGCAAGTAATTCGGTGGTGAAGCTTTTGATGAATGCAAG 543
QY      161 ArgArgAlaSerGluAlaArgGlyLysAlaLeuValIleTyrAsnGlnProTyrIleAsnTyr 180
Db      544 AGAAGGGCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 603
QY      181 SerArgThrValGlnTyrArgThrGlnGlyAlaValAlaGluAlaLysValGlyAlaLeu 200
Db      604 TCAAGAGGCTGCATACCAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 663
QY      201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
Db      664 GCATCTCTCATTCGATCCGTGGCTCTTTCATCTCAAGTCAAGTCTCAACAGGATATTAG 723

```

```

QY      221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db      724 GAATACACGAGATGCGCGGCCCAAGAAATTCACACCGCTGATTACGTTGAGAGATGCAAA 783
QY      241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMet-GlyAla 260
Db      784 ATGATGTCAGAAATGCGCTTCATGAGGATCAAAATTCATTCAGTTCAAGATGCGGCGC 843
QY      260 AlvsThrTyrProAspThrArgSerPheAsnThr-ValAlaGluIleThr-GlySerLys 279
Db      844 AAACACTACCCAGATAGTAATTCCTTACACACTGTAAGAGATCACTGGGAGCAAA 903
QY      280 TyrPro 281
Db      904 TATCCC 909
RESULT 8
BX360507      1126 bp      mRNA      linear      EST 05-MAY-2003
LOCUS
DEFINITION
BX360507 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSOD1071YA19 5-PRIME, mRNA sequence.
ACCESSION
BX360507
VERSION
BX360507.1 GI:30374439
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 1126)
AUTHORS
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSOD1071AA10Qp1cluster=4663.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitrogenCorporation1600
Faraday Avenue Genoscope sequence ID : CSOD1071AA10Qp1.
FEATURES
source
Location/Qualifiers
1..1126
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1071YA19"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      311 a      249 c      298 g      246 t      22 others
ORIGIN
Alignment Scores:
Pred. No.:      7.43e-141      Length:      1126
Score:          1312.50      Matches:      282
Percent Similarity: 94.02%      Conservative: 1
Best Local Similarity: 93.69%      Mismatches: 14
Query Match:    54.30%      Indels:      8
DB:            13      Gaps:      2
US-09-745-763-36 (1-472) x BX360507 (1-1126)
QY      1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGly 20
Db      210 ATGAATTCCTTATCTTCGATTTTTCGGTGTGTCACCTTTATCCCTGCTGCGG 269

```

21 LysAlaIleCysLysAsnGlyIleSerIysArgThrPheGluGluIleLeu 40  
 270 AAAGCTATATGCAAGATGGCATCTCTAGAGAGACTTTGAGAAATTAAGAAAGATA 329  
 41 AlaSerCysGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60  
 330 GCCACCTGTGAGATGTTGCTAAAGCAATCATCACTAGCTGTTTATGATGAAGCCAG 389  
 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
 390 AACAGATCTTATGACGATGGCACTTCTGTTGATGCTGTTGACCAAGATGAGTGGC 449  
 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100  
 450 TCCAGAACTTGAAGAAAGCCATCCAAATTATGTCACAAACCTGACAGAGATGGCTG 509  
 101 GluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluGluSerAla 120  
 510 GAGAAAGTTTCACTGAGGCACTGAGATATACCCCACTGGGAGAGGAGAAATCAGCT 569  
 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
 570 GTGATGCTGGAGCCAAAGATTCATAGATAGCATCTGGGCTTGGGACGACGACTGGG 629  
 141 ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160  
 630 ACTCTTCAGAAAGGCAATACAGCAAGATTTCTGGTGAGTCACTCTTTCATATACCTGAG 689  
 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180  
 690 AGAAGGCTCTGAGAAAGCAAGAGAGATGTTGTTTATTAACCAACTTACATCAACTAC 749  
 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaAlaLysValGlyAlaLeu 200  
 750 TCAGAGAGCGTGCAATACCGAAGCGAGGGGGGCGTGAAGCTGCCAAGGTTGGGCTTGG 809  
 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220  
 810 GCATCTTCATTCGATTCCTGGCTCTCTCTTCATCTACAGCTCCACAGAGATTTGAG 869  
 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGln 240  
 870 GAATACCAAGATGGGCTGGCCAAAGATTCACACAGCTGATTAACGGTGGAAAGATCGAGA 929  
 240 UmeMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
 930 AATGATGTCAGAAATGGCTTCTCATTTGGGATCAAAATGTCTCATTCAGCTAAAGATGGGG 989  
 260 ValValThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIle 280  
 990 S-AAAGACTACCCAGATCTGATCTCTTAC-ACGTGACCAAA-ATCAGCTGGAGACAA-W 1045  
 280 TyrProGluGlnValValLeuValSerGlyHisLeuAspSerTyrPheAspValGlyGlnGly 299  
 1046 ATCCAAAGAGCTACT---GKYAGTGGW---WCTGGAAATGGGATGTTGGGAGGGG 1098

RESULT 9  
 BG751497 902 bp mRNA linear EST 15-MAY-2001  
 LOCUS 60230076f1 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:4873770 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG751497  
 VERSION BG751497.1 GI:14062150  
 KEYWORDS EST.  
 ORGANISM Homo sapiens (human)  
 SOURCE Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 902)  
 AUTHORS NIH-MGC <http://mgi.mgi.nhl.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished

COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: [rsb@nih.gov](mailto:rsb@nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLW)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLW at:  
<http://image.lnl.gov>  
 Plate: L10M1752 row: 5 column: 19  
 High quality sequence stop: 817.  
 Location/Qualifiers  
 1..902  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4873770"  
 /rname="normal pigmented retinal epithelium"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 43"  
 /note="Organ: eye; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCAAGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

BASE COUNT 246 a 200 c 251 g 204 t 1 others

ALIGNMENT SCORES:  
 Pred. No.: 1,02e-139 Length: 902  
 Score: 1301.50 Matches: 279  
 Percent Similarity: 93.77% Conservative: 7  
 Best Local Similarity: 91.48% Mismatches: 6  
 Query Match: 53.85% Indels: 16-  
 DB: 10 Gaps: 3

US-09-745-763-36 (1-472) x BG751497 (1-902)

111 ProHisTyrGluArgGlyGluGluSerAlaValMetLeuGluProArgIleHisLysIle 130  
 2 CCCACCTGGAGAGAGGAGAAAGATCACTGTGATGCTGGAGCCAAAGATTCATTAAGATA 61  
 131 AlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIleThrAlaGluVal 150  
 62 GCCATCTGGGCTCTTGGGACGACGATGGGACTCTCCAGAAAGCATTAACGAGAAATT 121  
 151 LeuValValThrSerPheAspGluLeuGlnArgArgAlaSerGluAlaArgGlyLysIle 170  
 122 CTGGTGTGACCTCTTTCATGAACTGCAGAGAGGGCTTCAGAAAGGAGAAATG- 180  
 171 ValValTyrAsnGlnProTyrIleAsnTyrSerArgThrValGlnTyrArgThrGlnGly 190  
 181 GTTGTTATTAACCAACCTTACATCACTCAAGAGAGGTCGAATACCGAAGCAGAGGG 240  
 191 AlaValGluAlaAlaLysValGlyValAlaLeuAlaSerLeuIleArgSerValAlaSerPhe 210  
 241 GCGGTGAAGCTGGCAAGGTGGGCTTTGGCATCTCATTCATCCGTGGCTCTTC 300  
 211 SerIleTyrSerProHisThrGlyIleGlnGluTyrGlnAspGlyValProLysIlePro 230  
 301 TCCATCTACAGTCTCTCACACAGATTCAGGAATACCAAGATGGCGTCCCAAGATTTCA 360  
 231 ThrAlaCysIleThrValGluAspAlaGluMetSerArgMetAlaSerHisGlyIle 250  
 361 ACAAGCTGTATTAAGGTGAGAGATGCAAGATGATGTCAAAGATGGCTTCCATGGATC 420  
 251 LysIleValIleGlnLeuLysMetGlyAlaLysThrTyrProAspThrAspSerPheAsn 270  
 421 AAAATTGTCATTCAGCTAAAGATGGGCGCAAGAGCTTACCCAGATACGATTCCTTCAAC 480

QY 271 ThrValAlaGluIleThrGlySerIleTyProGluGlnValIleValSerGlyHis 290  
 DB 481 ACTGTAGACAGATCACTGGAGCAATATATCCAGAACAGGTGTGATGCTGTGATGACAT 540  
 QY 291 LeuAspSerTTPAPaValGlyGlnGlyAlaMet-AspAspGlyGlyGlyAlaPheIleSe 310  
 DB 541 CTGACACAGCTGGAGTGGGAGGAGGCTCCATGCGATATGCGGTGAGAGCTTTATATTC 600  
 QY 310 rTrpGluAlaLeuSerLeuIleIleYAspLeuGlyLeuArgProIleYAspGlyThrIleArgIle 330  
 DB 601 ATGGAGACACTCTCACTTATTAAGATCTTGCGCTGCTCCAAAGAGAGACTTGGGCT 660  
 QY 330 uValIleuTrpThr-AlaGluGluGlnGlyGlyValGlyAlaPheGlnTyTrpGlnLeuH 350  
 DB 661 GGTGCTGTGAGCTGGCAGAGAACAGAGGTGGAGTGGTGGTCCATGATTAATATGATTAC 720  
 QY 350 IeIleValAlaMetIleSerAsn-----TyrSerIleuValMetGluSerAspA 365  
 DB 721 ACAAGTAAATATTCACACTAACAGTGTGGGTGATGAGTCTT-----GAGC 768  
 QY 365 IaGlyThrPheLeu---ProThrGlyLeuGlnPheThr-GlySerGluIleValAlaArgAla 383  
 DB 769 CAGAACCTTTCTTACCACTGGCTGGCTGCAATTCACCTGGCAGTGAAGAGCCAGCC 828  
 QY 384 IleMetGluGluValMetSerIleuGlnProIleuAsnIleThrGlnValIleuSerHis 403  
 DB 829 ATCATGAGAGAGGATTATGAGCTGATGAAGCCT--CATATCACTCAGGTCCG-AAACAT 885  
 QY 404 GlyGluGly 406  
 DB 886 GGAAGAGGA 894

## RESULT 10

BX335995

LOCUS

BX335995 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA

clone CSOD1022YJ16 5-PRIME, mRNA sequence.

BX335995

BX335995.1 GI:30339459

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 919)

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 4663.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgj-bln/cluster.cgi?seq=CSOD1022DE080P1&amp;cluster=4663.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paradey Avenue Genoscope sequence ID : CSOD1022DE080P1.

## FEATURES

source

1..919

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CSOD1022YJ16"

/cissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

## BASE COUNT

273 a 206 c 226 g 208 t 6 others

## ORIGIN

## Alignment Scores:

Pred. No.: 2.25e-138

Score: 1290.00

Percent Similarity: 98.87%

Best Local Similarity: 98.87%

Query Match: 53.37%

DB: 13

Gaps: 0

US-09-745-763-36 (1-472) x BX335995 (1-919)

QY 1 MetIysPheLeuIlePheAlaPhePheGlyGlyValHisIleuLeuSerIleCysSerGly 20  
 DB 126 ATGAATTCCTTATCTTGCGATTTTGGGTGATTTTACCTTTATCCCTGTGCTGGG 185  
 QY 21 LysAlaIleCysIysAsnGlyIleSerIleArgThrPheGluGluIleLeuGluIle 40  
 DB 186 AAACATATGCAAGAAATGCGATCTTAARAGCACTTTGAAAGAAATTAARAGAAATA 245  
 QY 41 AlAserCysGlyAspValAlaIleValAlaIleIleAsnLeuAlaValTyrglyIleValGln 60  
 DB 246 GCCAGCTGTGGAGATGTTGCTAAAGCAATCAACCTAGCTGTTTATGTAAAGCCAG 305  
 QY 61 AsnArgSerTyrglyuArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
 DB 306 AACGATCTTATGAGCGATTTGGCACTTCTGTTGATCTGTGAGACCCAGCTAGTGC 365  
 QY 81 SerIysAsnLeuGluIleValAlaIleGlnIleMetTyrglnAsnLeuGlnIleAspGlyLeu 100  
 DB 366 TCCAAGAACTTAAAGAAAGCCATCAATTTATGACCAAACTGACAGAGATGGGCTG 425  
 QY 101 GluIysValHisIleuGluProValArgIleProHisIleProIleArgGlyGluIleSerAla 120  
 DB 426 GAGAAAGTTCACCTGGAGCCAGTGAATATCCCACTGGAGAGGAGGAGAAATACAGCT 485  
 QY 121 ValMetLeuGluProValGlyIleHisIleValIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
 DB 486 GTGATGCTGGAGCCAGAAATTCATTAAGATAGCCATCTCGGCTCTTGGCAGAGCATTCGG 545  
 QY 141 ThrProProGluGlyIleThrAlaGluValIleuValIleThrSerPheAspGluLeuGln 160  
 DB 546 ACTCTCCAGAAAGCATTAACAGCAGAAAGTCTGCTGGTGAACCTTTCAGATGAACCTGAG 605  
 QY 161 ArgArgAlaSerGluAlaArgGlyIleValIleValTyrglnGlnProTyrlIleAsnTyrl 180  
 DB 606 AGAAGGGCTTCAGAGAGAGGAGAGATTTGTTTATTAACCAACCTTATCATCACTAC 665  
 QY 181 SerArgThrValGlnTyrglyArgThrGlnGlyAlaValAlaGluAlaIleValIleValIleu 200  
 DB 666 TCAGAGAGGCTGCAATACCAAGCCAGAGGGGCGGTGAGAGCTCCAGAGTTGGGGCTTTG 725  
 QY 201 AlAserLeuIleArgSerValAlaIleSerPheSerIleTyrglnProHisIleGlyIleGln 220  
 DB 726 GCATCTCTCATTCATCGTGGCTCTCTTCCATTCACAGCTCTCCACACAGGATTTACG 785  
 QY 221 GluTyrglnAspGlyValProIleIleProThrAlaCysIleThrValGlnAspAlaGlu 240  
 DB 786 GAATTAACAGATGCGCGGCCCAAGATTCACACCTGTATTAAGGTGAAGAGACAA 845  
 QY 241 MetMetSerArgMetAlaSerHis-GlyIleIleValIleGlnIleuLeuMetGlyVal 260  
 DB 846 ATGATGTCAGAAATGGCTTCTCATGGGATCAAAATTTCTCATGCTTAARATGGGGGC 905  
 QY 260 alyeThrTyrlPro 264  
 DB 906 -AARACTTACCA 917

## RESULT 11

BUI46905

LOCUS

BUI46905 866 bp mRNA linear EST 03-SEP-2002

AGENCY: 7978204 Lupski dorsal root ganglion Homo sapiens CDNA

clone IMAGE:6185233 5', mRNA sequence.



ACCESSION BU146905  
 VERSION BU146905.1 GI:22660437  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 866)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished  
 JOURNAL  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Dr. James R. Lupski  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
 Plate: LLM13575 row: 1 column: 02  
 High quality sequence stop: 731.  
 Location/Qualifiers  
 1..866  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6185233"  
 /sex="male"  
 /tissue\_type="dorsal root ganglia"  
 /dev\_stage="adult, 36 yr"  
 /lab\_host="DH10B"  
 /clone\_lib="Lupski dorsal root ganglion"  
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site\_1: NotI; Site\_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCAACGCGTCG-3' and 5'-GACGTATCTAGATCGCGAGCGCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

BASE COUNT 257 a 182 c 233 g 191 t 3 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 6.62e-137 Length: 866  
 Score: 1277.00 Matches: 268  
 Percent Similarity: 92.54% Conservative: 5  
 Best Local Similarity: 90.85% Mismatches: 11  
 Query Match: 52.83% Indels: 11  
 DB: 13 Gaps: 4

US-09-745-763-36 (1-472) x BU146905 (1-866)

QY 23 IIECYLVSASNGIYIIESELYSARQTHR-PHEGIGIUIELVSGIUGIUIEALAE 42  
 DB 3 AATACCAAGATGGCATCTCTAGAGGACTTNTGAAAGAAAGAAAGAAATGCGCG 62  
 QY 42 ICYSGIYASPAVALALALYSAIAIIELESENLEUALAVALTYGTYLVALAGLNASAR 62  
 DB 63 CTGTGAGATGTCTGTAAGCAATCATCACTAGCTGTTTGTGTAAGCCCAAGACG 122  
 QY 62 GSEYTGIAUAGLEUALALEULEUVALASPTHYVALGTYPROAGLEUSERYSERIY 82  
 DB 123 ATCCATATGAGCATGCGACCTCTGCTGATACCTGTTGAGACCAAGACTAGGCTCCA 182  
 QY 82 SASNLEUGIUYSAIAIIEGNIEMERTYRGIPASNLEUGININASGTYLEUGIUY 102  
 DB 183 GAACCTAGAAAAGCAATCCAAATTAATGACAAAACCTGACGACAAAGATGGGCTGAGAA 242  
 QY 102 SVALHSLEUGIUPROVALARGILEPROHISTIRPGIUAUGIYGIUGIUSERALAVAME 122

DB 243 AGTTCACCTGAGGCAAGATATACCCACTGGGAGGGGAGAGAAAGATCAGCTGCAT 302  
 QY 122 TLEUGIUPROARGIIEHSILYSIIEAIAIIELEUGIYLEUGIYSESERIIEGIYTHPR 142  
 DB 303 GCTGAGGCAAGAAATTCATTAAGATAGCCATCTGGGTCTTGAGCAGCATTTGGACCTCC 362  
 QY 142 GPROGIUGIYIIEHRAIAGIUALLEUVALIYTHSERPHEASPGIULEUGIARGAR 162  
 DB 363 TCCAGAGGCAATTCAGCAGAAAGTTCTGTGTATATACCAACCTTACATCACTACTCAAG 422  
 QY 162 GAIASERIGUALAARGIYIIEVALIYIYASNGINPROTYRIEANTYSEAR 182  
 DB 423 GGCCCTCAGAAACAGAGGAGAGATTTGTTTATATACCAACCTTACATCACTACTCAAG 482  
 QY 182 GTHVALGINYTRGTHNGIYALAVAGIUALALAYVALGTYVALALASE 202  
 DB 483 GACGTCGAATACGAACGAGGGGGGGTGAAGCTGCCAAGGTGGGGCTTGGCATTC 542  
 QY 202 TLEUIEYSEYVALASERPHESERIEYTSERPROHISTIRGTYIIEGNGIUY 222  
 DB 543 TCTCATTCGATCCCTGGCTCTCTCTCATCTACAGCTTACACAGGTATTCAGAAATA 602  
 QY 222 TGINASPGIYALPROYSIIEPROTHALCYAIIETHRYVALGILUSPALAGIUMECME 242  
 DB 603 CCAGAGATGCGTGCCAAAGATTCACACAGCTGTATTAAGGTGAGAAATGCAAAATGAT 662  
 QY 242 TSEYRGMETALASERHSIGIYIIEVALIIEGNIENLEUSMETGTYALALYSTH 262  
 DB 663 GTCAAGATGGCTTCTCATYGGATCAAAATGTATTCAGCTAAAGATGGGGGCAAGAC 722  
 QY 262 RTYRPROAPHYRASPSPHEANTR-VALAIGIUIELTHGIYSEYRSTYR-PRO 281  
 DB 723 CTACCCAGATTCATCTTCTTCTTACACTGTCAGACAGATACCTGAGCAAAATATTC 782  
 QY 282 GIUGIUALLEUVALSER---GLYHSLEUASP---SEYTPASP---VALGYGIN 298  
 DB 783 GAACAGGGTTACTGGCTGAGTGACATCGAACACANCTGGAAATGTTGGGCGCA 842  
 QY 299 GLYALMETASPAEPIGYIYALAPHELESEYRTP 311  
 DB 843 GGGTGCCATGGGGA-----ATTGAATGG 866

RESULT 12  
 CB958693  
 LOCUS 739 bp mRNA linear EST 29-APR-2003  
 DEFINITION AGENCOURT 13778385 NIH-MGC 184 Homo sapiens cDNA clone  
 IMAGE:30351744 5', mRNA sequence.  
 CB958693  
 ACCESSION CB958693.1 GI:30214809  
 VERSION  
 KEYWORDS  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 739)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished  
 JOURNAL  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovics  
 cDNA Library Preparation: Clontech Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
 Plate: NDCM149 row: m column: 01  
 High quality sequence stop: 572.  
 Location/Qualifiers  
 1..739  
 /organism="Homo sapiens"



```

/mot_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30351744"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_184"
/notes="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site 1:
Site 1 (ggccatagcc); Site 2: Site 1 (ggccatagcc);
Library is oligo-dT primed and directionally cloned. cDNA
was prepared from a glandular pool of tissues from thyroid,
parathyroid, adrenal, cortex and pineal gland. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTATGACC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCCGAGCGCCGACATG-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.38
kb (range 0.60-3.5 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

BASE COUNT      192 a      171 c      195 g      181 t
ORIGIN

Alignment Scores:
Pred. No.:      3,18e-134      Length:      739
Score:          1253.00      Matches:      244
Percent Similarity: 99.59%      Conservative: 0
Best Local Similarity: 99.59%      Mismatches: 0
Query Match:    51.84%      Indels:      1
DB:             14      Gaps:      0

US-09-745-763-36 (1-472) x CB958693 (1-739)

Qy      197 ValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHis 216
Db      4 GTTGGGGGCTTTGGCATCTCTCATTCGATCCGCGCTCTTCCATCTACAGTCTTAC 63

Qy      217 ThrGlyIleGlnGlyTyrGlnAspGlyValProlylIleProthAlaCysIleThrVal 236
Db      64 ACAGGTATTCAGGAATACAGAGATGCGCTGCCCAAGATTCACACGCTGATTAACGCTG 123

Qy      237 GlnAspIleGlnMetSerMetSerArgMetAlaSerHisGlyIleValIleGlnLeu 256
Db      124 GAAATATGCAAGAAATGATGCAAGATGGCTTCTCTGAGATCAAAATGTCATTCAGCTA 183

Qy      257 LysMetGlyAlaValThrTyrProAspThrAspSerPheAsnThrValAlaGlnIleThr 276
Db      184 AAGATGGGGGCAAGACCTACCCAGATCTGATTCCTTCAACACTGTAGCAGATCTACT 243

Qy      277 GlySerIleTyrProGlnGlnValValLeuValSerGlyHisLeuAspSerTyrAspVal 296
Db      244 GGGAGCAAAATATCCAGAACAGATTGTACTGTCAGTGAATCTGACAGCTGGGATGTT 303

Qy      297 GlyGlnGlyAlaMetAspAspGlyGlyGlyAlaPheIleSerTyrGlnAlaLeuSerLeu 316
Db      304 GGGCAGGGGTGCATGATGATGCGCTGAGCTTATATATCATGGAGCACTCTCAGCTT 363

Qy      317 IleLysAspLeuGlyLeuArgProLysArgThrIleuArgLeuValLeuTyrThrAlaGln 336
Db      364 ATTAAAGATCTTGGCTGCTGCTCCAAAGAGACTCTGCGCTGGTGTCTGAGCTGCAAGA 423

Qy      337 GlnGlnGlyGlyValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAlaAsnIleSerAsn 356
Db      424 GAACAAGGTGGAGTGGTGGCTTCACGATATTATCATGACCAAGTAAATATTTCCACAC 483

Qy      357 TyrSerLeuValMetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThr 376
Db      484 TACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543

Qy      377 GlySerGlyLysAlaArgAlaIleMetGlnGlnValMetSerLeuLeuGlnProLeuAsn 396
Db      544 GGCAGTGAAGAGCCAGGCGCATCTGAGAAAGGTTATGAGCTTCTGAGCCCTTCAT 603

Qy      397 IleThrGlnValLeuSerHisGlyGlyGlyIleThrAspIleAsnPheTyrIleGlnAlaGln 416

```

```

Db      604 ATCACTCAGGTCTGACCATGAGGAAGGACAGACATCAACTTTTGATCCAGCTGA 663
Qy      417 ValProGlyAlaSerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisSerHis 436
Db      664 GTGCTTGAGCGCATCTTACTTGTATGATGATGATGATGATGATGATGATGATGATGAT 723

Qy      437 Gly-AspThrMet 440
Db      724 GGAAGACACCATG 736

RESULT 13
LOCUS      BK416896
DEFINITION BK416896 Homo sapiens PLACENTA Homo sapiens cDNA clone CSDB001YF04
5-PRIME, mRNA sequence.
ACCESSION  BK416896
VERSION     BK416896.1 GI:30650313
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 1012)
AUTHORS     Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished
COMMENT     Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSDB001YF04
Feng Liang Email: fliang@life.techn.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CSDB001YF04.

FEATURES
source
location/Qualifiers
1..1012
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDB001YF04"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

BASE COUNT      281 a      240 c      275 g      215 t
ORIGIN

Alignment Scores:
Pred. No.:      6.44e-134      Length:      1012
Score:          1253.00      Matches:      251
Percent Similarity: 99.21%      Conservative: 0
Best Local Similarity: 99.21%      Mismatches: 1
Query Match:    51.80%      Indels:      0
DB:             13      Gaps:      0

US-09-745-763-36 (1-472) x BK416896 (1-1012)

Qy      1 MetLysPheLeuIlePheAlaPheGlyGlyValHisLeuLeuSerLeuCysSerGly 20
Db      255 ATGAATCTCTTATCTTCGATTTTCGATGTTTCACCTTTATCTCTGCTGCGG 314

Qy      21 LysAlaIleCysLysAsnGlyLysSerLysArgThrPheGlnGlnIleGlnIle 40
Db      315 AAAGCTATATGCAAGAAATGCGATCTTGAAGGACTTTTGAAGAAATPAAAGAAATA 374

Qy      41 AlaSerCysGlyAspValAlaValAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60

```

```

Db 375 GCCACGTGGAGATGTTCTTAAGCAATCATCAACTGATGTTATGTAAGCCGAG 434
Qy 61 AenAserTyrGluArgLeuAlaLeuLeuValaSerThrValGlyProArgLeuSerGly 80
Db 435 AACCAATCTTAAGACGATGGACATTCCTGCTGATGATGATGAGACCCAGCTGATGAG 494
Qy 81 SerLeuAsnLeuGluValaAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100
Db 455 TCCAAAGAACCTTGAAGAAAGCCATCCAAATTAATGACCAAACTTCAGCAAGATGGCTG 554
Qy 101 GlnLysValaHisLeuGluProValaArgIleProHisTyrGlnArgGlyGlnIleSerAla 120
Db 555 GAGAAAGTTCACCTGGAGCCAGCTGAGAAATACCCCACTGGAGAGGGGAGAAATACAGCT 614
Qy 121 ValMetLeuGluProArgIleHisIleValaIleLeuGlyLeuGlySerSerIleGly 140
Db 615 GGATGCTGGAGCCAGAAATTCATAGATGACCATCTGGGCTTGGCAGCAGCATTTGGG 674
Qy 141 ThrProGluGluGlyIleThrAlaGluValaLeuValaIleThrSerPheAspGluLeuGln 160
Db 675 ACTCCTCCAGAGGCAATTCACAGCAAGTTCGTGGTGAACCTCTTTCATGATCACTGAG 734
Qy 161 ArgArgAlaSerGluAlaArgGlyLysIleValaIleTyrAsnGlnProTyrIleAsnTyr 180
Db 735 AGAAGGGCTCAGAAAGCAAGAGGAAAGATGTTGTTTATACCAACTTACATCAACTAC 794
Qy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValaGluAlaAlaLysValGlyAlaLe 200
Db 795 TCAGAGAGCGTGCAATACCGAAGCGAGGGGGCGGTGAGAGCTGCCAAGGTTGGGGCTTT 854
Qy 200 ValaSerLeuIleArgSerValaIleSerPheSerIleTyrSerProHisTyrGlyIleG 220
Db 855 GGCACTCTCATTCATCCGTGCTGCTCTCTTCCTTCATTCAGTCTTCACACAGGATTC 914
Qy 220 nGluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaG 240
Db 915 GGAATACAGCAGATGGGTGCCCAAGATTCCAACAGCTGTATTAAGTGGAAGATGCAAG 974
Qy 240 UMetSerArgMetAlaSerHisGlyIleLysIle 252
Db 975 AATGATGTCAGAAATGGCTTCTCATGAGATCAAAATT 1011

RESULT 14
LOCUS BX355939 996 bp mRNA linear EST 05-MAY-2003
DEFINITION BX355939 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1006Y115 3-PRIME, mRNA sequence.
ACCESSION BX355939
VERSION BX355939.1 GI:30382027
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Buxaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 996)
Full-length cDNA libraries and normalization
Unpublished
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr Web: www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4663.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1006AE08NPLcluster=4663.f. Contact:
Feng Liang Email: fliang@life.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS0D1006AE08NPL.
Location/Qualifiers
1..996
FEATURES
SOURCE

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1006Y115"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 270 a 231 c 223 g 269 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 5,31e-133 Length: 996
Score: 1244.00 Matches: 241
Percent Similarity: 99.59% Conservative: 0
Best Local Similarity: 99.59% Mismatches: 1
Query Match: 51.47% Indels: 0
DB: 13 Gaps: 0
US-09-745-763-36 (1-472) x BX355939 (1-996)
Qy 231 ThrAlaCysIleThrValGluAspAlaGluMetSerArgMetAlaSerHisGlyIle 250
Db 995 ACAGCTTATTTACGGTGGAAGATGCAAGAAATGATGTCAGAAATGGCTTCTCATGGGATC 936
Qy 251 LysIleValaIleGlnIleLysMetGlyValaLysThrTyrProAspThrAspSerPheAsn 270
Db 935 AAAATTGTCATTCAGCTAAAGATGGGGGCAAGACCTACCAATCTGATTCCTTCAAC 876
Qy 271 ThrValaIleGluIleThrGlySerLysTyrProGluGlnValaLeuValaSerGlyHis 290
Db 875 ACTGTAGCAGATCACTGGAGCAAAATATTCAGAAAGGTTGATCTGTCAGTGAGCAT 816
Qy 291 LeuAspSerTyrAspValaGlyGlnGlyAlaMetAspAspGlyGlyAlaPheIleSer 310
Db 815 CTGAGCAGCTGGATGTTGGGCAAGGTCATGATGATGGCGGTGAGCCTTATATATCA 756
Qy 311 ThrGluAlaLeuSerLeuIleLysAspLeuGlyLeuArgProLysArgThrLeuArgLeu 330
Db 755 TGGAGAGCACTCTCATTTAAAGATCTGGGCTGCGTCCAAAGAGACTTCGGCTG 696
Qy 331 ValLeuTyrThrAlaGluGluGlnGlyValaGlyAlaPheGlnTyrTyrGlnLeuHis 350
Db 695 GTGCTCTGCACTGCAGAGAAAGATGAGTGTGCTTCCAGTATATCACTTAC 636
Qy 351 LysValaAsnIleSerAsnTyrSerLeuValMetGluSerAspAlaGlyThrPheLeuPro 370
Db 635 AAGGTAAATATTTCCACTACAGCTGTGATGAGTCGACGACGAACTTCTTACC 576
Qy 371 ThrGlyLeuGlnPheThrGlySerGlyLysAlaArgAlaIleMetGluGluValaMetSer 390
Db 575 ACTGGGCTGCATATTCATCTGGCAGTGAAGAGCCAGGCCATCATGAGAGGAGTTATAGC 516
Qy 391 LeuLeuGlnProLeuAsnIleThrGlnValaLeuSerHisGlyGlnGlyThrAspIleAsn 410
Db 515 CTGCTGAGCCCTCATATCTACAGCTCTTGAGCAATGAGAGGACAGACATCAAC 456
Qy 411 PheTyrIleGlnAlaGlyValaProGlyAlaSerLeuAspAspLeuTyrLysTyrPhe 430
Db 455 TTTTGATTCAGAGGTGGAGTGCTCGAGCCAGTCTTACTTATGATCACTTAACTATTC 396
Qy 431 PhePheHisIleSerHisGlyAspThrMetThrValaMetAspProLysGlnMetAsnVal 450
Db 395 TTTCTTCATCACTCCACGAGACACATGATCTCTCATGTGATCCAAAGCATAAATGTT 336
Qy 451 AlaAlaAlaValaTyrPalValaValaSerTyrValaValaAspMetGluGluMetLeuPro 470
Db 335 GCTGCTGCTGTTGGGCTGTTGTTCTTATGTTGTTGACAGACATGAGAAATGCTGCT 276
Qy 471 ArgSer 472

```

Db 275 AGTCC 270

RESULT 15  
BO938234

LOCUS  
DEFINITION BO938234 926 bp mRNA linear EST 21-AUG-2002  
AGENCOURT 8932063 NIH\_MGC\_94 Mus musculus cDNA clone IMAGE:6465955  
5', mRNA sequence.

ACCESSION  
BO938234

VERSION  
BO938234.1 GI:2253712

KEYWORDS  
EST.

SOURCE  
Mus musculus (house mouse)

ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
1 (bases 1 to 926)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)

AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: The Cepko Laboratory  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML).  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LIML at:  
<http://image.llnl.gov>  
Plate: LLM13989 row: h column: 20  
High quality sequence start: 38  
High quality sequence stop: 605.  
Location/Qualifiers  
1..926  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6465955"  
/tissue="IMAGE:6465955"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_94"  
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 3.3 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC library."

BASE COUNT 252 a 200 c 259 g 215 t

ORIGIN

Alignment Scores:  
Pred. No.: 1.86e-129 Length: 926  
Score: 1213.00 Matches: 254  
Percent Similarity: 89.93% Conservative: 14  
Best Local Similarity: 85.23% Mismatches: 24  
Query Match: 50.19% Indels: 6  
DB: 13 Gaps: 0

US-09-745-763-36 (1-472) x BO938234 (1-926)

Qy 37 LyeGluGluIleAlaSerCysGlyAspValAlaIysAlaIleIleAsnLeuAlaValTyr 56  
Db 22 CGTGGGAAATAGCCAAATATGAAATGCTGCTTAAGCAATTATCACTGCTGTAT 81  
Qy 57 GilyysAlaGlnAsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyPro 76  
Db 82 GGTAAATACCAAGACCGGCTTATAGCGTTGGACCTTGTAGTGTGTTGGACCC 141  
Qy 77 ArgLeuSerGlySerLeuAsnLeuGluIysAlaIleGlnIleMetTyrGlnAsnLeuGln 96  
Db 142 AGACTGAGTGGCTTAAGAACTTAAGAAAGCTATTAATATATACCAAACTGCAAG 201  
Qy 97 GlnaArgGlyLeuGluIysValHisLeuGluProValArgGlyLeuProHisTyrGly 116  
Db 202 CAAGATGGGCTTGAAGAAATGTTCACTGGAGCAGTCAAGATACCCCACTGGGAGAGGGGA 261

Qy 117 GluGluSerAlaValMetLeuGluProArgIleHisIysIleAlaIleLeuGlyLeuGly 136  
Db 262 GAAGAATCTGAGAGATGCTTGAAGCCCGAATTCAAGATGCTATTCTAGAGCTTGGC 321  
Qy 137 SerSerIleGlyThrProProGluGlyIleThrAlaGluValLeuValIleThrSerPhe 156  
Db 322 AGCAGCATTTGGGACTCCCGCAGAGGAGCATCAAGACAGAGTGGTGGTGGCTCTTTT 381  
Qy 157 AspGluLeuGlnAAGAGAlaSerGluAlaArgGlyIysIleValIleTyrAsnGlnPro 176  
Db 382 GATGAACCTTCAAGAAAGACATTCAGAGCAAGAGGAGATCATTTTATTAACCACT 441  
Qy 177 TyrIleAsnTyrSerAcgThrValGlnTyrArgThrGlnGlyAlaValGluAlaIlys 196  
Db 442 TACACTGGCTATGAGAGACGTGTCATACCGGGTGAAGAGCTGGAGAGCTGCCAG 501  
Qy 197 ValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHis 216  
Db 502 GTGGGAGCTGGGACATCCATCATCAATCAATGAGTACCTTTTCACTACAGTCTCAC 561  
Qy 217 ThrGlyIleGlnGluTyrGlnAspGlyValProIysIleProThrAlaCysIleThrVal 236  
Db 562 ACAGGTATTCAGAAATATCAAGATGAGTGCCCAAGATTCACAGCTGTATCACAGTA 621  
Qy 237 GluAspAlaGluMetMetSerArgMetAlaSerHisGlyIleIysIleValIleGlnLeu 256  
Db 622 GAATGTCAGAAATGATGTCTCCGAATGGCTTCTGGGAGCAAAATGTCATTCACTG 681  
Qy 257 LysMetGlyAlaIysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThr 276  
Db 682 GAAATGGAGCAAGACCTATCCAGATGATGATCCCTCATACTGTTGACAGATCACT 741  
Qy 277 GlySerIysTyrProGluGlnValVal-LeuValSerGlyHisLeuAspSerTyr-Asp 296  
Db 742 GGGAGCATGTATCCAGAGAAAGTTGCTCGGTGACATTTGACAGCTGGGAGT 801  
Qy 296 AlGlyGlnGlyAlaMetAsp-AspGlyGlyValAla-PheIleSerTyrGlu-AlaLeu 315  
Db 802 TTGGGCAAGGTGCACTGGAATGAGGGGGGGGAACTTTTCATATCATGGGAAAGCACTCT 861  
Qy 315 er-LeuIleIysAspLeuGlyLeuArgProIysArgThrLeu 328  
Db 862 CCTTGGTTAAAGATCTTGGGGCTGGCCCCCAAGAGAGACTTC 903

Search completed: December 23, 2003, 01:44:34  
Job time : 2358 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: December 22, 2003, 23:17:31 ; Search time 95 Seconds  
(without alignments)  
2192.977 Million cell updates/sec

Title: US-09-745-763-36  
Perfect score: 2417  
Sequence: 1 MKKLIFAFPGVHLLSLCSG.....AVMAVSYVYADMEMLPRS 472

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODE=frame+.p2n.model -DEV=x1p  
-Q=/cgn2\_1/USPTO.spool\_p/US09745763/runat\_22122003\_113318\_25275/app\_query.fasta\_1.647  
-DB=Issued\_Patents\_NA -OPMT=faastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdd  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NOEM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09745763 -OCGN\_1\_1\_85@runat\_22122003\_113318\_25275 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2417	100.0	1863	4	US-09-482-273-28
2	1338	55.4	1134	4	US-09-482-273-95
3	229.5	9.5	1491	4	US-09-079-955-1
4	225.5	9.3	1640	3	US-09-330-095-2
5	182	7.5	1644	4	US-09-252-991A-2005
6	182	7.5	1755	4	US-09-252-991A-1721
7	166	6.9	2653	1	US-08-325-553-1
8	166	6.9	2653	2	US-08-394-152A-1
9	166	6.9	2653	4	US-08-705-477E-1
10	161	6.7	2133	4	US-09-164-034B-1
11	161	6.7	2387	4	US-08-705-477E-100
12	149	6.2	765	4	US-09-252-991A-1872

13	149	6.2	4403765	3	US-09-103-840A-2	Sequence 2, Appl1
14	149	6.2	4411529	3	US-09-103-840A-1	Sequence 1, Appl1
15	127	5.3	4080	4	US-09-016-434-1352	Sequence 1352, Ap
16	127	5.3	5020	3	US-09-161-244-1	Sequence 1, Appl1
17	126	5.2	780	1	US-08-325-553-27	Sequence 27, Appl
18	126	5.2	780	2	US-08-394-152A-27	Sequence 27, Appl
19	126	5.2	780	4	US-08-705-477E-27	Sequence 27, Appl
20	108.5	4.5	11935	4	US-09-634-238-401	Sequence 401, Ap
21	105	4.3	1230025	4	US-09-198-452A-1	Sequence 1, Appl1
22	103	4.3	7470	3	US-08-417-089-5	Sequence 5, Appl1
23	103	4.3	7470	3	US-08-695-651-5	Sequence 5, Appl1
24	103	4.3	7470	3	US-08-330-285-5	Sequence 5, Appl1
25	103	4.3	7470	3	US-08-695-421-5	Sequence 5, Appl1
26	103	4.3	7470	4	US-08-697-826A-9	Sequence 9, Appl1
27	102.5	4.2	1014	4	US-09-107-532A-2344	Sequence 2344, Ap
28	100.5	4.2	1419	1	US-08-242-098-39	Sequence 39, Appl
29	100.5	4.2	1974	4	US-09-423-439-59	Sequence 59, Appl
30	100.5	4.2	3177	4	US-09-252-991A-521	Sequence 521, App
31	100.5	4.2	80161	3	US-09-036-987A-1	Sequence 1, Appl1
32	100.5	4.2	80161	3	US-09-370-700-1	Sequence 1, Appl1
33	100.5	4.2	80161	4	US-09-603-207-1	Sequence 1, Appl1
34	100	4.1	36181	4	US-08-311-731A-120	Sequence 120, App
35	98.5	4.1	2805	4	US-09-252-991A-10208	Sequence 10208, A
36	98.5	4.1	4278	4	US-08-661-527-153	Sequence 153, App
37	98.5	4.1	124884	4	US-09-661-596A-76	Sequence 76, Appl
38	98	4.1	1847	4	US-09-321-017B-998	Sequence 998, App
39	97.5	4.0	1368	4	US-09-328-352-1713	Sequence 1713, Ap
40	97.5	4.0	1860	4	US-09-107-532A-1243	Sequence 1243, Ap
41	97.5	4.0	2048	3	US-08-776-251-1	Sequence 1, Appl1
42	97.5	4.0	2048	3	US-08-776-251-1	Sequence 1, Appl1
43	97	4.0	3300	1	US-08-194-290-6	Sequence 6, Appl1
44	97	4.0	3300	2	US-08-614-377A-6	Sequence 6, Appl1
45	97	4.0	3300	3	US-09-142-648B-6	Sequence 6, Appl1

## ALIGNMENTS

RESULT 1  
US-09-482-273-28  
; Sequence 28, Application US/09482273  
; Patent No. 6534631  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 71 Human Secreted Proteins  
; FILE REFERENCE: P2030P1  
; CURRENT APPLICATION NUMBER: US/09/482,273  
; EARLIER FILING DATE: 2000-01-13  
; EARLIER APPLICATION NUMBER: PCT/US99/15849  
; EARLIER FILING DATE: 1999-07-14  
; EARLIER APPLICATION NUMBER: 60/092,921  
; EARLIER FILING DATE: 1998-07-15  
; EARLIER APPLICATION NUMBER: 60/092,922  
; EARLIER FILING DATE: 1998-07-15  
; EARLIER APPLICATION NUMBER: 60/092,956  
; EARLIER FILING DATE: 1998-07-15  
; NUMBER OF SEQ ID NOS: 267  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 28  
; LENGTH: 1863  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-482-273-28  
Alignment Scores:  
Pred. No.: 8.56e-277  
Score: 2417.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 4  
Gaps: 0  
US-09-745-763-36 (1-472) x US-09-482-273-28 (1-1863)  
Length: 1863  
Matches: 472  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

```

Qy 1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCysSerGly 20
Db 99 ATGAAATTCCTTATCTTCGCAATTTTCGGTGGTGTTCACCTTTATCCCTGGCTGG 158
Qy 21 LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleGlyGluIle 40
Db 159 AAACCTATATCAAGAAATGGCATCTCTAAGAGAGACTTTTGAAGAAATGAAGAAATA 218
Qy 41 AlaSerCysGlyAspValAlaLysAlaIleIleLeuLeuAlaValTyrGlyLysAlaGln 60
Db 219 GCCAGCTGTGAGAGATGTGTCAAGCAATCAATCACTAGCTGTTATGTGAAGCCAG 278
Qy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80
Db 279 AACGATCTTATGAGCATTTGGCACTTCTGGTGTATCTGTGGACCAAGCTAGATGGC 338
Qy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100
Db 339 TCCAAGAACTTGAAGAAAGCCATCCAAATTTGTACAAACCTGCACAGATGGGCTG 398
Qy 101 GluLysValHisLeuGluProValArgIleProHisTyrProGluArgGlyGluGlnSerAla 120
Db 399 GAGAAAGTTCACTGTGAGCCAGTGAATACCCCACTGGAGAGGGAGAAAGAAATCAGCT 458
Qy 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140
Db 459 GTGATGCTGGAGCCAGAAATTCATAGATAGCAATCCCTGGGCTTGGAGCAGCATGGG 518
Qy 141 ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluGln 160
Db 519 ACTCTCCAGAAAGGATTAACAGAGAGTCTGGGTGGAGCACTTTTGATGAATCAGTGG 578
Qy 161 ArgArgAlaSerGluAlaArgGlyLysValLeuValTyrAsnGlnProTyrIleAspTyr 180
Db 579 AGAAGGGCTCAAGAAAGCAAGAGAGAGATGTGTGTTATTAACCAACTTACATCACTAC 638
Qy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaLysValGlyAlaLeu 200
Db 639 TCAAGAGAGGTGCATACCAAGACGAGGGGGCGGTGGAAGCTGCCAAGGTTGGGCTTG 698
Qy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
Db 699 GCATCTCTCATTCGATCCGTGGCTCTTCTTCATCTCAAGCTCAACAGATTCAG 758
Qy 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysGlyIleThrValGluAspAlaGln 240
Db 759 GAATACCGAGATGGGTGCCCAAGATTCCAACAGCTGTATTAACGTGGAAGATGCAGAA 818
Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db 819 ATGATGTCAAGATATGGCTTCTCATATGGATCAAAAATGTTCATTCAGCTAAAGATGGGCA 878
Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280
Db 879 AAGACCTACCAAGATACGATTCCTTCAACACTGTAAGAGATCATCTGGAGCAAAATAT 938
Qy 281 ProGluGlnValIleLeuValSerGlyHisLeuAspSerTyrAspValGlyGlnGlyAla 300
Db 939 CCAAGAACGGTTGATCTGTGACGTGACATCTGCACCTGGAGATGGTGGAGGGTCC 998
Qy 301 MetAspAspGlyGlyAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320
Db 999 ATGATGATGCGCGGTGAGCTTATATCATATGGAAGACATCTCATTTTAAAGATCTT 1058
Qy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTyrThrAlaGluGluGlnGly 340
Db 1059 GGGGCTGTCCAAAGAGAGACTCTGGGCTGGGTCTGCACTGCAGAGAAACAAGTGGGA 1118
Qy 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAlaGlnIleSerAsnTyrSerLeuVal 360
Db 1119 GTTGGTGCCTTCCAGTATATATCAAGTACAAAGGTAATATTTTCCAACTTACAGTCTGGTG 1178
Qy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLys 380

```

```

Db 1179 ATGAGTCTGACGAGAACCTTCTTACCCAGCTGGCTGCATATTCATCGCAGTGAAG 1238
Qy 381 ALAARGALAIleMetGluGluValMetSerLeuLeuGlnProLeuAlaIleThrGlnVal 400
Db 1239 GCCAGGGCCATCATGAGAGAGGATTAAGCTGTGTGAGCCCTCCAAATATCACTCAGTGC 1298
Qy 401 LeuSerHisGlyGluGlyThrAspIleAsnPheThrIleGlnAlaGlyValProGlyAla 420
Db 1299 CTGAGCCATGAGAGAGGACACATCATCTTTTGATTCAGCTGAGATGCCAGACC 1358
Qy 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440
Db 1359 AGTCTACTGATGACTTATACAGATATTTCTTCCATCATCTCCACGAGACACCATG 1418
Qy 441 ThrValMetAspProLysGlnMetAsnValAlaAlaValTyrProAlaValAspTyr 460
Db 1419 ACTGTCAATGATCCAAAGACAGATGAATGTGTGCTGTGTGGGTGTTGTTTCTTAT 1478
Qy 461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db 1479 GTTGTTCAGACATGAGAAAGATTCCTCTAGTTC 1514

RESULT 2
US-09-482-273-95
: Sequence 95, Application US/09482273
: Patent No. 6534631
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: 71 Human Secreted Proteins
: FILE REFERENCE: P2030P1
: CURRENT APPLICATION NUMBER: US/09/482, 273
: EARLIER FILING DATE: 2000-01-13
: EARLIER APPLICATION NUMBER: PCT/US99/15849
: EARLIER FILING DATE: 1999-07-14
: EARLIER APPLICATION NUMBER: 60/092,921
: EARLIER FILING DATE: 1999-07-15
: EARLIER APPLICATION NUMBER: 60/092,922
: EARLIER FILING DATE: 1998-07-15
: EARLIER APPLICATION NUMBER: 60/092,956
: EARLIER FILING DATE: 1998-07-15
: NUMBER OF SEQ ID NOS: 267
: SOFTWARE: Patent Ver. 2.0
: SEQ ID NO 95
: LENGTH: 1134
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-482-273-95

Alignment Scores:
Pred. No.: 5,07e-149 Length: 1134
Score: 1338.00 Matches: 261
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 55.36% Indels: 0
Gaps: 0

US-09-745-763-36 (1-472) x US-09-482-273-95 (1-1134)
Qy 211 SerIleTyrSerProHisThrGlyIleGlnGluTyrGlnAspGlyValProLysIlePro 230
Db 1 TCCATCTACAGTCTCTCACAGGATTCAGGAATACCAAGATGGCGGCCAAGATTCGA 60
Qy 231 ThrAlaCysIleThrValGluAspAlaGluMetCysArgMetAlaSerHisGlyIle 250
Db 61 ACAACCTGTATTAAGGTGGAAGATGACGAATGATGTCAAGAAATGAGGCTTCTCATGGGATC 120
Qy 251 LysIleValIleGlnLeuLysMetGlyAlaLysThrTyrProAspThrAspSerPheAsn 270
Db 121 AAAATGTTCATCACTGAAGATGGGGCAAGACCTTACCAAGATCTGATTCCTTCAAC 180
Qy 271 ThrValAlaGluIleThrGlySerLysTyrProGluGlnValLeuValSerGlyHis 290

```

```

Db      181 ACTGACAGAGATCACTGGAGCAATATCCAGAACAGTTGACTGTCAGTGACAT 240
Qy      291 LeuAspSerTrpAspValGlyGlnGlyAlaMetAspAspGlyGlyValAlaPheIleSer 310
Db      241 CTGGACAGCTGGAGTGTGGCCAGGTGTCATGATGATGAGGAGGCTTTATATCA 300
Qy      311 TTPGluAlaLeuSerLeuIleLysAspLeuGlyLeuArgProLysArgTrpLeuArgLeu 330
Db      301 TGGAGACACTCTCATTTAATTAAGATCTTGGGCTGGCTGCAAGAGAGACTCTGGCTG 360
Qy      331 ValLeuTrpThrAlaGluGlnGlnGlyGlyValAlaPheGlnTrpTrpGlnLeuHis 350
Db      361 GTGCTCTGAGCTGCAAGAAACAAGGTGAGTTGCTGCTTCCAGTTATTCAGTTACAC 420
Qy      351 LysValAlaSerLeuSerLeuTrpSerLeuValMetGluSerAspAlaGlyTrpPheLeuPro 370
Db      421 AAGGTAATATTTTCAACTACAGTCTGTATGAGAGTCTGACGAGAACTCTTACCC 480
Qy      371 ThrGlyLeuGlnPheThrGlySerGlyLysValAlaArgAlaIleMetGluGluValMetSer 390
Db      481 ACTGGCTGCATTCATCTGCGAGTGAAGAAAGCCAGGCG-ATCATGAGAGAGTTATGAC 539
Qy      391 LeuLeuGlnProLeuAsnIleThrGlnValLeuSerHisGlyGlnGlyThrAspIleAsn 410
Db      540 CTGCTGACAGCCCTCAATATCATCTCAGTCTCTGAGCCATGAGAAAGGACAGACATCAAC 599
Qy      411 PheTrpIleGlnAlaGlyValProGlyAlaSerLeuAspAspLeuTrpLysTrpPhe 430
Db      600 TTTTGGATCCAAAGTGAAGTCCCTGAGAGCCAGTCTACTGATGACTTATATCAATATTTTC 659
Qy      431 PhePheHisSerHisSerHisGlyAspThrMetThrValMetAspProLysGlnMetAsnVal 450
Db      660 TTCTTCATCACTCTCCACGAGACACATGACTGTCATGATCCAAAGCAGATGAAGTT 719
Qy      451 AlaAlaAlaValTrpAlaValAlaSerTrpValValAlaAspMetGluGluMetLeuPro 470
Db      720 GCTGCTCTGTTGGGCTGTGTTCATTATGTGTGACATGAGAAATGCTGCT 779
Qy      471 ArgSer 472
Db      780 AGGTCC 785

```

## RESULT 3

```

US-09-079-955-1
; Sequence 1, Application US/09079955A
; Patent No. 6465209
; GENERAL INFORMATION:
; APPLICANT: Alexander Blinkovsky
; APPLICANT: Kimberly Brown
; APPLICANT: Elizabeth Gollighly
; APPLICANT: Tony Byun
; APPLICANT: Thomas Mathiasen
; APPLICANT: Lene V. Kolof
; APPLICANT: Chigusa Shizuka
; APPLICANT: Chigusa Shizuka
; TITLE OF INVENTION: Methods For Producing Protein
; TITLE OF INVENTION: Hydrolyses
; FILE REFERENCE: 5253.500-US
; CURRENT APPLICATION NUMBER: US/09/079.955A
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Aspergillus oryzae
US-09-079-955-1

```

```

Alignment Scores:
Pred. No.: 3.39e-17
Score: 229.50
Percent Similarity: 42.12%
Best Local Similarity: 22.59%
Length: 1491
Matches: 96
Conservative: 83
Mismatch: 175

```

```

Query Match: 9.50% Indels: 71
DB: 4 Gaps: 19
US-09-745-763-36 (1-472) x US-09-079-955-1 (1-1491)
Qy      77 ArgLeuSerGlySerLysAsnLeuGluValAlaIleGlnIleMetTrpGlnAsnLeuGln 96
Db      157 CGGCTCTTGTGTGTAAGCCACAGCAGACAGCTTAATCATCTTACGAGAGCTGAAG 216
Qy      97 GlnAspGlyLeuGluLysValHisLeuGluProValArgIleProHisTrpGluArgGly 116
Db      217 AAGACTGCTACTATGATGATGCTTCAAGACAGCTT-CAAGTGAAGCTGTGAGCAATGCC 273
Qy      117 GlnGluSerAlaValMetLeuGluProArgIleHisIleValAlaIleLeuGlyLeuGly 136
Db      274 GACCAGAGCTCAAGTGGGCGGATGAGAAATGAC-----GCCAAG 315
Qy      137 SerSerIleGlyThrProGluGlnGlyIleThrAlaGluValAlaValTrpSerPhe 156
Db      316 ACCATGACCTAAGTCCACAGCTGAGTGCACCGCGATGACCGCTGCAAGAACCTG 375
Qy      157 AspGluLeuGlnArg--ArgAlaSerGluAlaArgGlyLysIleValAlaIleTrpAsnGln 175
Db      376 GATGACGACGAGCGGATTAACCATCGATGCGAGGCAAGGTGCGCTGATCAAGCT 435
Qy      176 -----ProTrpIleAsnTrpSerArgThrValGlnTrpArgThrGlnGlyAlaVal 192
Db      436 GAGAAATGCCCTTCCGCGCACAGTCC-----GTT 465
Qy      193 GlnAlaAlaLysValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPheSerIle 212
Db      466 CTGCTGCTCAAGGCAAGCGCGCGGCTTGATTTGATTAACAATGCGCGGATTCATG 525
Qy      213 TyrSerProHisThrGlyIleGlnGluTrpGlnAspGlyValProLysIleProThrAla 232
Db      526 GCGGACACCTTGAGCGCGGCGAG-----AGTATAGAGACCGGATTCGCCATTGTC 579
Qy      233 CysIleThrValGlnAspAlaGluMetMetSerArgMetAlaSerHisGlyIleLysIle 252
Db      580 GGTATCAGCTTGAGTGAAGTGGCCAGAACCTGATCAAGCTTCTGAGGTGAG--TCGGTA 636
Qy      253 ValIleGlnLeuLysMetGlyAlaLysTrpTrpAspThrAspSerPheAsnTrpVal 272
Db      637 TCTGTGATCTGTGGGCGTGAATGATGAAGCAGAGAACCTGAG--ACGTATTAAGTTGTC 693
Qy      273 AlaGluIleThrGlySerLysTrpTrpGluGlnValAlaLeuValSerGlyHisLeuAsp 292
Db      694 GCGCAGACGAAAGGCGCGCAT--CCGAAACAAGCTCTCCGCTGCGTGGCCACACGAG 750
Qy      293 SerTrpAspValGlyGlnGlyAlaMetAspAspGlyGlyGlyAlaPheIleSerTrpGlu 312
Db      751 TCAGTCAAGCGCGGCGCTGTGATACACAGATGCTGGGCAATTATTAACACTTGGTC 810
Qy      313 AlaLeuSerLeuIleLysAspLeuGlyLeuArgProLysArgTrpLeuArgLeuValLeu 332
Db      811 ATTGCCAAGCGCTCAGCAGACTCTCGTC-----AAGATGCCGCGCTTCTCTTC 864
Qy      333 TrpThrAlaGluGlnGlnGlyGlyValAlaPheGlnTrpTrpGlnLeuHis 347
Db      865 TGGACACAGAGGAGTTCGCTGCTGGGCGAGC--AACTACTACGTTCCCATCTGAAT 921
Qy      348 -----GlnLeuHisLysValAlaIle-----SerAsn 356
Db      922 GCCACGAGCTGAACCAAGTCCAGCTTACTGACTGACATTCGACATGATGCGCTCACTAAC 981
Qy      357 TyrSerLeu--ValMetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPhe 375
Db      982 TAGCCCTCATGATCTATGAAGTGAATGATGATGCGCGCTTCAACAGAGGAGCG- 1035
Qy      376 ThrGlySerGlyLysValAlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeu 395
Db      1036 GCCGTTCCGCCGATCGAGAAACTGTCGAGAC-----TACTACGACTCCATC 1086

```

Db 541 GGAAGATGTGGCTTCGGC-----GACAAGTGGCTTCGGCTGCC 579  
 Oy 196 LysValGlyAlaLeuAlaSerLeuLeuArgSerValAlaSerPheSerIleTyrSerPro 215  
 Db 580 AAAGCCCAAGGCGCGCGCTTCGATTCTTAATACAAATGTGCAGAGATCATAGGAGCACC 639  
 Oy 216 HisThrGlyIleGlnGlnIuTyrGlnAspGlyValProlysIleProThrAlaCysIleThr 235  
 Db 640 CTTCGGCGCGCGCAGACTGACAGAGACCGTAATTCCGCCATT-----GTCCGATATAGC 693  
 Oy 236 ValGluAspAlaGluMetMetSerArgMetAlaSerHisGlyIleLysIleValIleGln 255  
 Db 694 TTGGAGAGATGGCCAGAGAGCTGATCATCACTGCTGAGCTGGA---TCGGATCTGTGAT 750  
 Oy 256 LeuLysMetGlyAlaLysThrTyrProAspThrAspSerPheAsnThrValAlaGluIle 275  
 Db 751 CTGTGGGTGTGATGCACAGAGAGAACCGTACG---ACGTATTAACGTTATTCGGCGACAGC 807  
 Oy 276 ThrGlySerIleTyrProGluGlnValValLeuValSerGlyHisLeuAspSerTyrAsp 295  
 Db 808 AAGGCGCGCGCAT---CCGAACAAATGTGCTGCGCGCTGGTGAGCCCACTGACTCGCTGAG 864  
 Oy 296 ValGlyGlnGlyAlaMetAspAspGlyGlyValAlaPheIleSerTyrGluAlaLeuSer 315  
 Db 865 GCGGCGCTCGGTATCAATGACAGATGCTCGGGCATTTTACGAAACCGTGCTGGCCAA 924  
 Oy 316 LeuIleLysAspLeuGlyLeuArgProLysArgThrLeuArgLeuValLeuTyrThrAla 335  
 Db 925 GCGCTGACGCGAGTACTCCGTC---AAGAATGCCGTGCGCTTCTTCTTGACGGCC 978  
 Oy 336 GluGluGlnGlnGlyValAlaGlyAlaPheGlnTyrTyr-----Gln 348  
 Db 979 GAGAGATTGCGTCTCCGCGGACG---AACTACTGACTCTCCCATCTGAAATGCCACCGAG 1035  
 Oy 349 LeuHisIleValAsnIle-----SerAsnTyrSerLeu 359  
 Db 1036 CTGAACAAAGATCAGACTGTACTGTACCTGACATCATATGCCCTTCGCCCACTAGCCCTC 1095  
 Oy 360 ---ValMetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySer 378  
 Db 1096 ATGATCTATGACGGGTGACGATGCGCGCTTCAACACAGACGGACG---GCCGAGTCC 1149  
 Oy 379 GluLysAlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThr 398  
 Db 1150 GCCCAAGATCGAAGAAAGCTTCGAGAGC-----TACTACGACTCCATCGACTTGCTT 1200  
 Oy 399 GlnValLeuSerHisGlyGlnGly---ThrAspIleAsnPheThrPleGlnAlaGlyVal 417  
 Db 1201 CATATCCGACCCAGCTTCGACGAGACGTTCCGATTACAGAGCCTTTATCTGAAGGCAATT 1260  
 Oy 418 ProGlyAlaSerLeuLeuAspLeu----- 426  
 Db 1261 CCGGCGGTGAGACTCTTCACGGGCGCGAGGCGCATATGTCCGAAGAAACGAAGCGT 1320  
 Oy 427 -----TyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440  
 Db 1321 TGGGAGAGTCAAGCGCGCGCTGACCTACAGACCCAACTACACACGCGCGGAGACCAACATG 1380  
 Oy 441 ThrValMetAspProLysGln-----MetAsnValAlaIleAlaValTyrAlaValVal 458  
 Db 1381 ACCAACTCAACCATGAAGCCTTCGTGATCATCCAAAGCACACAGCCTTCGCGCTGCC 1440  
 Oy 459 SerTyr 460  
 Db 1441 ACCTAC 1446

RESULT 5  
 US-09-252-991A-2005/c  
 ; Sequence 2005, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOKOMAS



```

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2005
; LENGTH: 1644
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2005

Alignment Scores:
Pred. No.: 1,74e-11      Length: 1644
Score: 182.00           Matches: 95
Percent Similarity: 38.71%      Conservative: 61
Best Local Similarity: 23.57%    Mismatches: 141
Query Match: 7.53%             Indels: 106
DB:                               Gaps: 19

US-09-745-763-36 (1-472) x US-09-252-991A-2005 (1-1644)

QY 65 GluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGlySerIysAsnLeu 84
DB 1395 GAGACATCGCCACTCTCAAGAC-----GGCAACGGCGCGCGCGCCAGCCGGGCTAC 1342
QY 85 GluLysAlaIleGlnIleMetTYrGlnAsnLeuGlnIleAspGlyLeuGluValHis 104
DB 1341 CAGGCTCTCCGCTGACTGATGACAGACCTCTGCAAGAACCGGCTAC---AAGGTACAGC 1285
QY 105 LeuGluProValaLargIleProHisTrpGluArgGlyGluGluSerAlaValMetLeuGlu 124
DB 1284 GTGAGCGCTTCCTCCGTTACCGCTAC-----TAC 1255
QY 125 ProArgIleHisIysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProGlu 144
DB 1254 CCGAAA-----GGCGGGGTAGCTGAGACGGCAGCGCTGCGCAG 1216
QY 145 -----Gly 145
DB 1215 CCGGTACCTACGATGAGAGAGAAAGACTTCACTACTGTGCGACGACGAGCGAGCGAC 1156
QY 146 IleThrAlaGluValIleValVal-----ThSer 155
DB 1155 GTCAACGGCCAGAGGTGTCCTCGGTGACCTGTCTCGCGCGCGCAACACTCCACCGACG 1096
QY 156 PheAspGluLeuGlnArgAlaSerGluAlaArgGlyIysIleValIleValIleValIle 175
DB 1095 GGTTCGAGAGGAGAAAGACTTCCCACTTCCCGCGCGCTGATCGCGCTATCCAGCGC 1036
QY 176 ProTYrIleAsnTYrSerArgThrValGlnTYrArgThrGlnGlyAlaValGluAla 195
DB 1035 GGCACCTGCACCTTCGAG-----CAGAAGCGCGAGAACGCCGCG 997
QY 196 LysValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPheSerIleTYrSerPro 215
DB 996 GCCCGCGCGCGCGCGGTGATCATCTTCAACAGCGGCAAC-----ACCGACGAC 946
QY 216 HisThrGlyIle-----GlnGluTYrGlnAspGlyValProIleIle 229
DB 945 CGCAAGGCGCTCGAAGACGTACCGGTGGCGAGTCTTCAGAGGGCGGATCCCGGTATC 886
QY 230 ProThrAlaCysIleThrValGluAspAlaGluMetSerArgMetAlaSerHisGly 249
DB 885 -----TTGCGCACCTACGACACAGCGCGTGGCTGTGCGACAGACCCGAGCCTGAG 835
QY 250 IleLysIleValIleGlnIleLeuMetClyAlaIleTYrThrProAspThrAspSerPhe 269
DB 834 TTGACACTGTGTGTGACGTG-----GTACGCAAGAAAGACCGAGACCTTAC 790

```

```

QY 270 AsnThrValAlaGluIleThrGlySerIysTYrProGluGlnValIleValSerGly 289
DB 789 AACGTGTCCCGGAG---ACCGCGCGGCAACCCGAAACAGTGTGATGTGGCGCG 733
QY 290 HisLeuAspSerTrpAspValGlyGlnGlyAlaMetAspAspGlyGlyGlyAlaPheIle 309
DB 732 CACTTCACCTCGGTGTGCAAGGCCCGGTATCAACAGACGTTGGCGAGCGCGCC 673
QY 310 SerTrpGlu---AlaLeuSerLeuIleLysAspLeuGlyLeuArgProLysArgThrLeu 328
DB 672 CAATGAGATGCGCGTGTCTGCTGGCCAGAGCGCTG-----CCGTCACACAGGTG 622
QY 329 ArgLeuValLeuTrpThrAlaGluGlnGlyValGlyValAlaPheGlnTYrGln 348
DB 621 CGTTGCGCTGTGTGGCGCGCGGAGAACCGCGCTGTGGCTGACCCACTATGTCAG 562
QY 349 LeuHisLysValAsnIleSerAsnTYrSerLeuValMetGluSerAspAlaGlyThrPhe 368
DB 561 -----AACCTCGCCCGGAGAG----- 544
QY 369 LeuProThrGlyLeuGlnPheThrGlySerGluLysAlaArgAlaIleMetGluGluVal 388
DB 543 -----AAGAAAGATCAAGGCTTACTGAC----- 517
QY 389 MetSerLeuLeuGlnProLeuAsnIleThrGlnValLeuSerHisGlyGlnGlyThrAsp 408
DB 516 TTGACATGATCGGCTCGCGCACTTGCAGACTTATATATAGACGCGAGCGTTCCGAC 457
QY 409 IleAsnPheTrpIleGlnAlaGlyValProGlyAlaSerLeuLeuAspAspLeuTYrLys 428
DB 456 -----TTCCGCTTCAG---GGTCCCGCGCGCTCGCGCCCATCGAGCCCTGTTGAA 406
QY 429 TyrPhePhe 431
DB 405 GCCTACTTC 397

RESULT 6
US-09-252-991A-1721
; Sequence 1721, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1721
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1721

Alignment Scores:
Pred. No.: 1.93e-11      Length: 1755
Score: 182.00           Matches: 95
Percent Similarity: 38.71%      Conservative: 61
Best Local Similarity: 23.57%    Mismatches: 141
Query Match: 7.53%             Indels: 106
DB:                               Gaps: 19

US-09-745-763-36 (1-472) x US-09-252-991A-1721 (1-1755)

QY 65 GluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGlySerIysAsnLeu 84
DB 367 GAGACATCGCCACTCTCAAGAC-----GGCAACGGCGCGCGCGCCAGCCGGGCTAC 420
QY 85 GluLysAlaIleGlnIleMetTYrGlnAsnLeuGlnIleAspGlyLeuGluValHis 104

```

```

Db      421 CAGGCTCCGTCAGTACGTGAGACACCTGCGAGAAAGCGGCTAC---AAGGTACGC 477
      105 LeuGIuPProValAlaGleProHisTrpGluArgGlyGluGluSerAlaValMetLeuGlu 124
      478 GTGAGCCCTTCCCTTCACCGCTAC-----TAC 507
      125 ProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGlu 144
      508 CCGAAA-----GACCGGGTACCTGAGCGCCACCTGCGCGAG 546
      145 -----Gly 145
      547 CCGGTACCTACGAATGGAGAGAGACTTACACTGCTGCGACAGCCGAGCAGCAGC 606
      146 IleThrAlaGluValLeuValAla-----ThSer 155
      607 GTACCCGCGAAGGTGTCCTCGGTGACCTGTCGCGCGGACACACTCCACCGC 666
      156 PheAspGluLeuGluArgAlaSerGluAlaArgGlyLysIleValValTyraGln 175
      667 GGTTCGAGGCGGAGAGACTTCCCACTTCCGCGCGCTGATGCGCTGATCCAGCGC 726
      176 ProTyrIleAsnTyrSerArgThrValGlnTyrArgThrGlnGlyAlaValAlaAla 195
      727 GGCACCTGCACCTTCGAG-----CAGAAGCGCGAGACGCCGCG 765
      196 LysValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerPro 215
      766 GCCCGCGCGCGCGCGGTATCATCTTCAACCGCGGCAC-----ACCGACGAC 816
      216 HisThrGlyIle-----GlnGluTyrGlnAspGlyValProLysIle 229
      817 CGCAAGGCGCTGGAACGTACCTGCGCGAGCTCTACGAGGCGGACCTCCGCTGATC 876
      230 ProThrAlaCysIleThrValGluAspAlaGluMetMetSerArgMetAlaSerHisGly 249
      877 -----TTGCGCACCTAGACAAACGCGGTGCTGTCGACACCCCGGACCTGAG 927
      250 IleLysIleValIleGlnLeuLysMetGlyAlaLysThrTyrProAspThrAspSerPhe 269
      928 TTGCACTGTGCTGTCGACGTG-----GTACGCAAGAACACCGACCTAC 972
      270 AsnThrValAlaGluIleThrGlySerLysTyrProGluGlnValValLeuValSerGly 289
      973 AACGTGTGCGGAG---ACCGTGTGCGGACCCGAAACACGTGTATGTGTCGCGCG 1029
      290 HisLeuAspSerTrpAspValGlyGlnGlyAlaMetAspAspGlyGlyAlaPheIle 309
      1030 CACCTCGACTGCGGTGTGGAAGGCCCGGTATCAACGACACCGTTCGGGACGCCGCC 1089
      310 SerTrpGlu---AlaLeuSerLeuIleLysAspLeuGlyLeuArgProLysArgThrLeu 328
      1090 CAACGTGAGATGCGCTGCTGTGCGCAAGCGCTG-----CCGTCACAAAGGTG 1140
      329 ArgLeuValLeuTrpThrAlaGluGlnGlyValGlyAlaPheGlnTyrTrpGln 348
      1141 CGCTTCGCTGTGCGGCGCGGAGAGCGGCTGTGCTGACACCACTACCTGCGAG 1200
      349 LeuHisLysValAsnIleSerAsnTyrSerLeuValMetGluSerAspAlaGlyThrPhe 368
      1201 -----AACCTCGCCCGGAAAGAG----- 1218
      369 LeuProThrGlyLeuGlnPheThrGlySerGluLysAlaArgAlaIleMetGluGluVal 388
      1219 -----AAGAAGAAAGATCAAGGCTTACTCTACCTAAC----- 1245
      389 MetSerLeuLeuGlnProLeuAsnIleThrGlnValLeuSerHisGlyGlnGlyThrAsp 408
      1246 TTCGACATGATCGCTCGCCGAACTTCGCAACTTCATCTATGACGGGACGCTTCGAC 1305
      409 IleAsnPheTrpIleGlnAlaGlyValProGlyAlaSerLeuLeuAspLeuTyrLys 428

```

```

Db      1306 -----TTGCGCTTCAG---GTCGCGCGGCTGCGCCGCAATCGAGCGCTGTTCGAA 1356
      429 TyrPhePhe 431
      1357 GCCTACTTC 1365

RESULT 7
US-08-325-553-1
Sequence 1, Application US/08325553
Patent No. 5538866
GENERAL INFORMATION:
APPLICANT: Israel, Ron S.
APPLICANT: Heston, Warren D.W.
APPLICANT: Fair, William R.
TITLE OF INVENTION: THE PROSTATE-SPECIFIC MEMBRANE ANTIGEN
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESS: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/325,553
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/973,337A
FILING DATE: 05 NOV 1992
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 1747/41426
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)-977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2653 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Carcinoma
IMMEDIATE SOURCE:
CLONE: Prostate-Specific Membrane Antigen
FEATURE:
NAME/KEY: CDS
LOCATION: 262..2511
US-08-325-553-1

Alignment Scores:
Pred. No.: 3,03e-09 Length: 2653
Score: 166.00 Matches: 105
Percent Similarity: 32.47% Conservative: 70
Best Local Similarity: 19.48% Mismatches: 169
Query Match: 6.87% Indels: 195
DB: 1 Gaps: 23

US-09-745-763-36 (1-472) x US-08-325-553-1 (1-2653)
3 PheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuLeuCysSerGlyLysAla 22

```

```

Db      355  TTCTTTCTCGGCTTCTCTTGGGCTTTTAAATCTCCAAATGAGCTACTAAC 414
Qy      23  11eCysLysaenGlyIleSerLysArgThrPheGluGluIleLysGluGluIleAlaSer 42
Db      415  ATTACTCCAAAGCATTAATATGAAAGCATTTTGGATGATGAAAGCTGAG----- 465
Qy      43  CysGlyAspValAlaIleIleAsnLeuAlaValTyrGlyLysAlaGlnAsnArg 62
Db      466  -----AACATCAAGAAAGTTCTTATATATTT----- 492
Qy      63  SerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProAsnGlyLeuSerGlySerLys 82
Db      493  -----ACACAGATCACCACTTTAGCAGAAACAGAA 522
Qy      83  -----AsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAsnArgGly 99
Db      523  CAAAACTTTCAGCTTGCAAGCAAAATTCATCCAGTGAAAGAAATTTGGCTGGATTTCT 582
Qy      100  LeuGluLysValHisLeuGluProVal-----ArgIleProHisTyr 113
Db      583  GTTAGCTAGCAGCATTTATGATGCTGCTGTTGCTCAACCAATAGACTCATCCCACTAC 642
Qy      114  -----GluArgGlyGluGlu-----SerAlaValMetLeuGluProArg 126
Db      643  ATCTCAATTAATTAAAGAAATGAGAAATGATTTTCAACACATCATTTATTGAAACA--- 699
Qy      127  IleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIle 146
Db      700  -----CCCTCTCCAGGATAT 714
Qy      147  ThrIleGluValLeuValIleThrSerPheAspGluLeuGlnArgAlaSerGluAla 166
Db      715  GAAATGTTTCGATATTTGACACCTTCAGTCTTCTCTCTCTCAAGAAATG---CCA 771
Qy      167  ArgGlyLysIleValIleValTyrAsnGlnProTyrIleAsnTyrSerArgThr----- 183
Db      772  GAGGCGCATCTAGTG-----TATGTTAACTATGACAGAACTGAAAGACTTC 816
Qy      184  -----ValGlnTyrArgThrGlnGlyAlaValGluAlaIleVal 197
Db      817  TTTAAATTTGAAAGGAGCATGAAATCAATTCCTCTGGAAATTTGAAATTTCCAGATAT 876
Qy      198  Gly-----AlaLeuAlaSerLeuIleArgSerVal 207
Db      877  GGGAAAGTTTCAAGAGAAATTAAGTTAAATATGCCAGCTGCAAGGGCCAAAGAGATC 936
Qy      208  AlaSerPheSer-----IleTyrSerProHisThrGlyIleGlnGluTyrGln 223
Db      937  ATTCTTACTCCGACCTGCTGACTGACTTGTCTCT-----GGGGTGAAGTCTTATCCA 990
Qy      224  Asp----- 224
Db      991  GATGTTTGAATCTTCTCGAGGTGTGTCCAGCGTGAAATATCTAAATCTGAATGCT 1050
Qy      224  ----- 224
Db      1051  GCAGAGAACCTTTCACACCAAGTTTACCAGCAATGATATGCTTATAGCGCTGAATT 1110
Qy      225  -----GlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db      1111  GCAGAGCGTGTGTCTTCCAAAGATTTCTGTTCAATTTGATGATTAATGATGACAG 1170
Qy      241  MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db      1171  AAGCTCTTGA--- 1194
Qy      261  LysThrTyrProAsp----- 265
Db      1195  TCAGACCAACCAAGATGAGCTGAGAGAGAAATCTCAAAAGTCCCTTACATGTTGACCT 1254
Qy      265  ----- 265

```

```

Db      1255  GGCTTACTGGAACCTTTTCTACCAAAAGTCAAGATGACATCTCACTTACCAATGAA 1314
Qy      266  ---ThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyrProGluGlnVal 284
Db      1315  GTGACAGAAATTTTACATTTGATAGTACTCTCAGAGGACAGCTGAAACCAAGATAT 1374
Qy      285  ValLeuValSerGlyHisLeuAspSerTyrPaspValGlyGlnGlyAlaMetAspArgGly 304
Db      1375  GTATTCTGGAGAGTCAACCGGACTCATAGGCTGTT-----GGTGTATTTGACCTCAG 1428
Qy      305  GlyGlyAlaPheIleSerTyrGlu-----AlaLeuSerLeuIleLysAspLeuGly 321
Db      1429  AGTGAGACACTGTGTTCTCAATGAAATTTGTGAGAGAGCTTTGGAACACTGAAAGGAG 1488
Qy      322  LeuArgProLysArgThrLeuArgLeuValLeuTyrThrAlaGluGluGlnGlyVal 341
Db      1489  TGGAGACTTGAAAGAAATTTGTTTGGCAAGCTGGATGCAAGAAATTTGGCTTCTT 1548
Qy      342  GlyAlaPheGlnTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuValMet 361
Db      1549  GGTTCCTCTAGTGGGAGAG-----GAGAAATCAAGACTTCTTCAA 1590
Qy      362  GluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyVal 381
Db      1591  GAGCGTGGCGTG---GCTTATATTAAATGCTGACTCATCTTATAGAGAAATCACTCTG 1647
Qy      382  ArgAlaIleMetGluGluValMet---SerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db      1648  AGAGTTGATTGTACACCGCTGATGACAGCTTGATAC-----AACCTTAAACAAAGAG 1701
Qy      401  Leu---SerHisGlyGluGlyThrAsp-IleAsnPheTyrIleGlnAlaGlyVal 417
Db      1702  CTGAAAGCCCTGATGAGAGCTTTGAAAGCAATCTTATGAAAGTTGAGACTA 1756

```

## RESULT 8

US-08-394-152A-1

Sequence 1, Application US/08394152A

Patent No. 5935818

GENERAL INFORMATION:

APPLICANT: Israeli, Ron S.

APPLICANT: Heaton, Warren D.W.

APPLICANT: Falt, William R.

TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND

NUMBER OF SEQUENCES: 48

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Cooper &amp; Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM 330 466 DX2

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/394,152A

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 41426-B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2653 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

```

?      TOPOLOGY: linear
?      MOLECULE TYPE: CDNA
?      HYPOTHETICAL: NO
?      ANTI-SENSE: NO
?      ORIGINAL SOURCE:
?      ORGANISM: Homo sapiens
?      TISSUE TYPE: Carcinoma
?      IMMEDIATE SOURCE:
?      CLON: Prostate-Specific Membrane Antigen
?      FEATURE:
?      NAME/KEY: CDS
?      LOCATION: 262..2511
US-08-394-152A-1

```

```

OY      224 ----- 224
Db      1051 GCAGAGAACCTCTCACACACAGGTTACCAGCAAAATGAATATGCTTATAGCGGTGAATT 1110
OY      225 -----GlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db      1111 GCAGAGCGTGTGGTCTTCCAAAGTAATCTGTTTCATCCATTTGGATACTATGATGACAG 1170
OY      241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db      1171 AAGCTCCCTAGAA-----AAAATGGGTGC 1194
OY      261 LysThrTyrProAsp----- 265
Db      1195 TCAGACACCACCAATGATACAGCTGGAGAGAAAGTCTCAAAGTCCCTCAATGTGGACCT 1254
OY      265 ----- 265
Db      1255 GCGTTTACTGGAACCTTTTCTACACAAAAGTCAAAGTGCATCATCCACTTACCAATGAA 1314
OY      266 ---ThrAspSerPheAsnThrValAlaGluIleThrLysSerLysTyrProGluGlnVal 284
Db      1315 GTGCACAAGAAATTTCAAATGTGATATGTACTCTCAGAGAGACAGTGAACCAAGACAGATAT 1374
OY      285 ValLeuValSerGlyHisLeuAspSerTyrAspValGlyGlnGlyIleMetAspAspGly 304
Db      1375 GTCAATTCTGGAGAGTCAACCGGACTCATGGGTGTTT-----GGTGGTATTGACCCCTCAG 1428
OY      305 GlyIleValPheIleSerTyrGlu-----AlaLeuSerLeuIleLysAspLeuGly 321
Db      1429 AGTGAGACAGCGTGTGTTTCATGAAATTTGTGAGAGAGCTTTGAAACACTGAAAAAAGGAGGG 1488
OY      322 LeuArgProLysAspGlyThrLeuAspGluValLeuThrPheLysGluGlnGlyVal 341
Db      1489 TGGAGACCTGAGAAACAAATTTGTTTGCAGAGCTGGAGTCCAGAAAGAAATTTGGTCTTCTT 1548
OY      342 GlyAlaPheGlnIleTyrGlnLeuHisLeuValAsnIleSerAsnTyrSerLeuValMet 361
Db      1549 GGTCTCTAGTAGTGGGACAGG-----GAGATTTCAGAGCTCCTTCAC 1590
OY      362 GluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLysAla 381
Db      1591 GAGCGTGGCGCGT---GCTTATATTAAAGCTGACTCATCTATAGAGGAAACTACACTCTG 1647
OY      382 ArgAlaIleMetGluGluValMet---SerLeuGlnProLeuAsnIleThrGlnVal 400
Db      1648 AGAGTTGATGTTGACACCGCTGATGTACAGCTGTGTACAC-----AACCTAACAAAGAG 1701
OY      401 Leu---SerHisGlyGluGlyThrAsp-IleAsnPheTyrIleGlnAlaGlyVal 417
Db      1702 CTGAAAAGCCCTGATGAAAGCTTTGAAAGGCAAACTCTTATGAAAGTTGACATA 1756

RESULT 10
US-09-164-034B-1
GENERAL INFORMATION:
APPLICANT: Mincheff, Milcho S.
            Loukinov, I. Dmitri
            Zoubak, Serguei
TITLE OF INVENTION: Immunotherapy of Cancer Through Expression
                    of Truncated Tumor- or Tumor-Associated Antigen
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: William S. Ramsey,
            Ramsey, Cook, Looper & Kurlander, LLC
            STREET: 10420 Little Patuxent Parkway, Suite 250
            CITY: Columbia
            STATE: Maryland
            COUNTRY: USA
            ZIP: 21044
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: PC

```

```

OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/164,034B
FILING DATE: 30-Sep-1998
ATTORNEY/AGENT INFORMATION:
NAME: Ramsey, William S.
REGISTRATION NUMBER: 32,715
REFERENCE/DOCKET NUMBER: br11
TELECOMMUNICATION INFORMATION:
TELEPHONE: (410) 992-9660
TELEFAX: (410) 992-9540
SEQUENCE DESCRIPTION: SEQ ID NO: 1
US-09-164-034B-1

Alignment Scores:
Pred. No.: 8,3e-09 Length: 2133
Score: 161.00 Matches: 93
Percent Similarity: 33.48% Conservative: 63
Best Local Similarity: 19.96% Mismatches: 139
Query Match: 6.66% Indels: 171
Gaps: 21

US-09-745-763-36 (1-472) x US-09-164-034B-1 (1-2133)
QY 76 ProAglSerGlySerIys-----AenLeuGluValAlaIleGlnIleMetIyr 92
DB 115 CCACATTTCAGCAGACAGACAAACTTCAGCTTCGAAACCAATTCATCCACAGTGG 174
QY 93 GlnAenLeuGlnIleApsIylLeuGluValIleLeuGluProVal-----108
DB 175 AAGAATTTCGCTGATTCGTGAGCTAGACATTAATGATGCTCTGTCCTACCA 234
QY 109 -----ArgIleProHisItr-----GluArgIleGluIu-----Ser 119
DB 235 AATAAGACTCATCCCAACTACATCTCAATTAATTAATGAATGAATGAAATTTTCAAC 294
QY 120 AlaValMetLeuGluProArgIleHisIleValIleLeuGluLeuGluSerSerIle 139
DB 295 ACATCATATTTCGAAACA-----312
QY 140 GlyItrProProGluGluIleThrAlaGluValLeuValIleThrSerPheApsIleu 159
DB 313 -----CCTCCCTCCAGATATGAAATGTTTCGATATTCACACCTTCAGTCTTTC 366
QY 160 GlnAgtArgAlaSerGluAlaArgGlyIleValIleValIleThrAsnGlnProIleAsn 179
DB 367 TCTCTCTCAAGAAATG---CCAGAGGGCGATCTAATG-----TATGTTAAC 408
QY 180 TyrSerArgThr-----ValGlnItrArgThrGlnGly 190
DB 409 TATGCAACCAACTGAAGACTCTTTAAATTGGAACGGGACATATAATCAATTCCTGCG 468
QY 191 AlaValGluAlaAlaIleValIleGly-----AlaLeu 200
DB 469 AAAATTGTAATGTCAGATATGGAAGATTTTCAGAGAAATTAAGTTAAAAATCCAG 528
QY 201 AlaSerLeuIleArgSerValAlaSerPheSer-----IleTyrSerProHis 216
DB 529 CTGGAGGGGGCCAAAGAGATCTTCATCTCCAGCCCTGCTGACTACTTGTGCTCT---585
QY 217 ThrGlyIleGlnGluTyrGlnAps-----224
DB 586 ---GGGGTGAAGTCTTATTCAGATGCTTGGAATCTTCTCGAGGTGTGTCCAGCGTGA 642
QY 224 -----224
DB 643 AATATCTTAATTCGAATGTCGAGAGACCTCTCACACAGCTTACCCAGCAATGA 702
QY 225 -----GlyValProValIleProThrAlaCys 233
DB 703 TATGCTTATAGCGGTGAATTCAGAGGCGTGTGCTTCCTCCAAAGATTCTGTTTCATCA 762

```

```

QY 234 IleThrValGluApsAlaGluMetSerArgMetAlaSerHisGlyIleValIleVal 253
DB 763 ATGGATATCTATGATGACAGAACTCTCTGAA-----795
QY 254 IleGlnLeuIleMetGlyAlaIleValIleThrTyrProAsp-----265
DB 796 -----AAAATGGGTGGCTCAGACACACAGATAGCAGCTGAGAGGAAGTCTCAAA 846
QY 265 -----265
DB 847 GTGCCCTCAATGTTGACCTGCTTACTGGAACCTTTCTACACAAAAGTCAAGATG 906
QY 266 -----ThrAspSerPheAsnThrValAlaGluIleThrGly 277
DB 907 CACATCCACTCTACCAATGAAGTGAAGATTAATTAATGATGATGATGATGATGATGATG 966
QY 278 SerIysItrProGluGlnValIleValIleValIleSerGlyHisLeuApsSerItrPheValGly 297
DB 967 GCAGTGGAAACCAACAGATATGTCATTCCTGGAGGTGACCGGAGCTCATGGGTGTT---1023
QY 298 GlnGlyValMetAspApsGlyGlyValAlaPheIleSerItrGlu-----AlaLeu 314
DB 1024 ---GGTGATTCAGCTCTCAGAGTGGAGCAGCTGTTCTTATGAAATTTGAGAGAGCTTT 1080
QY 315 SerLeuIleIysApsLeuGlyLeuArgProIysArgThrLeuArgLeuValIleTyrThr 334
DB 1081 GCAGACTGAAAGAAAGAGGTGAGACCTTAGAAGCAATTTGTTTGCAGCTGGGAT 1140
QY 335 AlaGluGlnGlnGlyValIleValAlaPheGlnItrTyrGlnLeuHisIleValAsnIle 354
DB 1141 GCAGAAAGAAATTTGCTCTTCCTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1182
QY 355 SerAsnItrSerIleValMetGluSerAspAlaGlyIleThrPheLeuProThrGlyLeuGln 374
DB 1183 GAGATTCAGACCTCTTCAAGACGTCGCTG---CCTTATATTAATGCTGACTATCT 1239
QY 375 PheThrGlySerGluValAlaArgAlaIleMetGluIleValMet-----SerLeuLeuGln 393
DB 1240 ATAGAGGAACACTACCTGAGAGTGAATGTTGACCGCGATGATGATGATGATGATGATG 1299
QY 394 ProLeuAsnIleThrGlnValLeu---SerHisGlyGlnGlyIleThrAsp-IleAsnPheTr 412
DB 1300 -----AACCTAAACAAAGAGCTGAAAGCCCTGATGAGAGGCTTTGAAGCAATCTCT 1353
QY 412 PileGlnAlaGlyVal 417
DB 1354 TATGAAAGTTGACTA 1369

RESULT 11
US-08-705-477E-100
; Sequence 100, Application us/08705477E
; Patent No. 6569432
; GENERAL INFORMATION:
; APPLICANT: Israeli, Ron S.
; APPLICANT: Heaton, Warren D.W.
; APPLICANT: Fair, William R.
; APPLICANT: Overfield, Ouathex
; APPLICANT: Pinto, Don
; TITLE OF INVENTION: PROSTATE-SPECIFIC MEMBRANE ANTIGEN AND USES THEREOF
; FILE REFERENCE: 1769/41426-G
; CURRENT APPLICATION NUMBER: US/08/705,477E
; CURRENT FILING DATE: 1996-08-29
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 100
; LENGTH: 2387
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-705-477E-100

Alignment Scores:
Pred. No.: 9.99e-09 Length: 2387
Score: 161.00 Matches: 93

```



```

QY      196  LySValGAlAAlenulaSerLeuIleArgSeValAlaserPheSerPro 215
Db      242  GCGCCGGCGCGCGCGGGGTGATCATCTTCAACAGGCACC-----ACGCACGAC 292
QY      216  HisrNglyIle-----GIInglTyrgInaApjValProlysile 229
Db      293  CGCAAGGCGCTGGAGAAGCACCGTGCGGCGAGTCCTCAGAGGGCGCATCCGGTGATC 352
QY      230  ProThrAlaCyellEthrValGluaBpaIagImetwecSerArgMetalaSerHisgly 249
Db      353  -----TTGCCCACTTACGACAACCGCGTGGCTGTGTGGAGACCCCGACTGECAG 403
QY      250  ILeysIlleValIleGlnLeuLysmetGlyAlalyeThyrTrpRoasPrThraSerPhe 269
Db      404  TTGCACCTGTGTGTGCACGTG-----GTACGCAGAAGAACCGAGACTTAC 448
QY      270  AantHrValAlaGluIleThrygLyserTyTrpGluInValIleValaIseryIly 289
Db      449  AACGTGTGCGCGAG---ACCGTGCGCGCACCCGAACAACGTGTGTGTGTGGCGG 505
QY      290  HisLeuAspSerTrpAspValGlyGlnGlyAlamechaspApolyGlyAlaPheIle 309
Db      506  CACCTCGACTCGGTGTTCGAAGCCCCGGGTATCAAGACACAGCTTCGGGCACGCGCC 565
QY      310  SerTrpGlu---AlaLeuSerLeuIleYasPleuglyLeuakRppPolySatgThrLeu 328
Db      566  CAACGTGAAGATGCGCTGTCTGTGGCCAAAGCGCTG-----CCGTCAACAAGTG 616
QY      329  ArgLeuValleuTrpThraIagiIngInglyIyValaIyAlaphegIntyTrygin 348
Db      617  CGCTTCGCTGTGGGGGGCGCGAGAACCGCGCTGTGTGGCTGTCAACCCACTACGTGCAG 676
QY      349  LeuHisLeuValAsnIleSer-----AerTyR 357
Db      677  -----AACCTGCGCCGGAAGAGAAGATCAAAGCCTTACCTGAACCTTC 724
QY      358  SerLeuValMezCluSerAspAlaGlyThrPheLeu 369
Db      725  GACATCATCGCGCTCGCGCAACTTCGGCACTTATC 760

RESULT 13
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-2007.00
; CURRENT APPLICATION NUMBER: US/09103, 840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 0.0615 Length: 4403765
Score: 149.00 Matches: 97
Percent Similarity: 34.73% Conservative: 60
Best Local Similarity: 21.46% Mismatches: 181
Query Match: 6.16% Indels: 114
```

DB: 3 Gap: 21

US-09-745-763-36 (1-472) x US-09-103-840A-2 (1-4403765)

QY 75 G1yPrOATgLeuSerG1ySerLySAsnLeuGluValAla1leGln1leMetYrGlnAen 94

Db 505153 GGCACtCGCGCGGTGGGCAcCCCTGGATATCAAGGCACGGCTGCATATGTGGTAAACACA 505212

QY 95 LeuGln1naSpG1yLeuGln1ySValHisLeuGluProValArg1leProHisTrpGlu 114

Db 505213 CTGCGCAACAGCGGCTTTGAT-----GTCCAAACCCCGGAGTCTCC 505254

QY 115 -----ArgG1yGluGluSerAlaVal1leMetLeuGluProArg1leHisLyS1le 130

Db 505255 GCTCGCGGTTCACAGCGCAAAAGGCGGTGGTGGACCTTCGGCGGC-----AACACCGTG 505308

QY 131 Ala1leLeuG1yLeuG1ySerSer1leG1yThrProProGluG1y1leThrAlaGluVal 150

Db 505309 GAGCGGAGGGCGGCTGAGTACAGCCTCGGCACACCGCGGAGGGGTGAAGCGGCGGCTG 505368

QY 151 LeuValVal1ThrSerPheAspG1yLeuGln1naArgAlaSerGlu----- 165

Db 505369 GTGGCTGCCCGCCGCGAGCAGACAGTCCGGGCTGCAGCTCGCGAGTACAGACAGGCTGCCG 505428

QY 166 AlaArgG1yLyS1leValVal1yTrAsnGlnProYr1leAsnTySerArg1yThrValGln 185

Db 505429 GTGTCCGGTGGCGGTGTCTGGTAATCGCGGCGCTGTGCTTTGGC----- 505476

QY 186 TyrArg1yGlnG1yAlaValGluAlaAlaLySValG1yAlaLeuAlaSerLeu1leArg 205

Db 505477 -----CAGAGGAAGACCGCACGCCGCGGAGCGGCTGGCGCTGATCATTTGCT 505527

QY 206 -----SerValAlaSerPheSer1leYrSerProHisThrG1y1leGlnGlu 221

Db 505528 GACACATCGACGAGCAGCGCGATGGAGCGGCACTCGGGGGCTTAATACGACGCTC----- 505581

QY 222 TyrGlnaSpG1yValProLyS1leProThrAlaCyS1leThrValGluAspAlaGluMet 241

Db 505582 -----AAGTATCCCGGTGTGAGTGTACCAAG----- 505608

QY 242 MetSerArgMetAlaSerHisG1y1leLyS1leVal1leGlnLeuLyS----- 257

Db 505609 -----TCGGTCGATTCACGTTACGACACCGGACAGTCTGGG 505641

QY 258 ---MetG1yAlaLyS1yThrTyRProAsp1yThrAspSerPhe-----Asn1yThValAla 273

Db 505642 CCAACACCGCTCAAGCTCAcCGCGGAGCAGCACCAGTTCAGAGGCCGCGCAACGTCATCGCG 505701

QY 274 Glu1le---ThrG1ySer1yS1yTyRProGluGlnValValLeuValSerG1yHisLeuAsp 292

Db 505702 CAGAGGAAGACGGGTGTCTGCG-----GCCAACGTGGTATGATGACAGGTGCGGCTTTGGAC 505755

QY 293 Ser1yTrpAspValG1yGlnG1yAlaMetAspAspAspG1yG1yAlaPhe1leSerTrpGlu 312

Db 505756 AGCGTTCCGGAAGAGACCGGACATCAACGACACGGCTCGGAGTGGCT-----GCG 505806

QY 313 AlaLeuSerLeu1leLySAspLeuG1yLeuArgPro-----LySArg1yThrLeuArgLeu 330

Db 505807 GTTTCGAAAGCGGCGAGTGCAGCTGGGGAACTCACCGGCAATGTGTCCAAACCGGTAACGGTTC 505866

QY 331 ValLeuTrp1yThrAlaGluGlnG1yGlnG1yValG1yAlaPheGln1yTyRTrpGlnLeuHis 350

Db 505867 GCCTTCTCGGGGCGCGAGGAATTCGGCTGATGGGTGACGAACACTACGTCGAGTGCCTG 505928

QY 351 LysValaAn-----1leSerAsnTySerLeuValMetGluSer 363

Db 505927 GACATCGACGCGCTCAAAAGCATCGCGCTGTATCTGAACCTTGACACATGTGTGGCGTCCCG 505986

QY 364 AspAlaG1y1yThrPheLeuProThrG1y-----LeuGln1yThrG1ySer 378

Db 505987 AACCCGGGTACTTCACTACGATACAGGTACACAGTGCCTGCCGTAGACCCCGCGGTACG 506046

QY 379 -----GluLyS1leAlaArgAla1leMetGluGluValMetSerLeuLeuGln1yPro 394



```

Db 506047 CCGGTGTGCCCCGAAGGCTCCGGCGGTATCGAGCGCAAGTTCGTGCGCTTATCGTAAGTGG 506106
Qy 395 Leuamnllethr---GlnValLeuSerHisglYleuGluYThrAspIleAsnPhetripIle 413
Db 506107 GCCGGCAAGACCGCGGAGACACCTTCGTTCGACGGTCGGTCCGACTACGACGAGCGCTTCACG 506166
Qy 414 GlnAlaGlyValProGlyAlaSerLeu----- 422
Db 506167 CTGGCGGGTATCCCTTCGGGTGGCTGTTCCTCGGCGGTAGTCAGAAAGTCCGCCGAG 506226
Qy 423 -----LeuAspAspLeuYrYleYrPhe 430
Db 506227 CAAGCCAGCTCTGGGGCGGACCGCGACGAGCTTCGATCCCAACTAT----- 506277
Qy 431 PheHeHisSerHisglYAspThrMetThrValMetAspProlysglMet----- 448
Db 506278 -----CACCAAGAACAGACACCTCGACCATATCGACCGACCGCGGTATC 506328
Qy 449 AsnValAlaAlaValAlaValAlaValAlaValSerYr 460
Db 506329 AACGGCGCTGGCGCTCGCTACCGGTGGTTGTAT 506364

RESULT 14
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 0.0617 Length: 4411529
Score: 149.00 Matches: 97
Percent Similarity: 34.73% Conservative: 60
Best Local Similarity: 21.46% Mismatches: 181
Query Match: 6.16% Indels: 114
DB: 3 Gaps: 21

US-09-745-763-36 (1-472) x US-09-103-840A-1 (1-4411529)
Qy 75 GlyProAlaGlyLeuSerGlySerLeuAsnLeuGluYValAlaIleGlnIleMetYrGlnAsn 94
Db 503710 GGCACTCCCGCGGTGGGACCCCTGCTATCGACGGCAGCTGCTATGCTGTAACACA 503769
Qy 95 LeuGlnIleAspGlyLeuGluYValHisLeuGluProValArgIleProHisIleArgIle 114
Db 503770 CTGGCGGACACGCGTTTGTAT-----GTGCAAAACCCCGAGTTCTCC 503811
Qy 115 -----ArgGlyGlnGluSerAlaValMetLeuGluProArgIleHisIleVal 130
Db 503812 GCTCGCGTTCACAGCGCGAAAGGGGTGTGACCTCGCGGCG-----AACACCGTG 503865
Qy 131 AlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIleThrAlaGluVal 150
Db 503866 GAGCGAGAGCGCGCTCGAGTACACCTCGGACACCGCGGAGGTGACGCGCGCGCTG 503925
Qy 151 LeuValAlaThrSerPheAspIleuGluIleArgAlaSerGlu----- 165

```

```

Db 503926 GTGGCTGCCCCCGCGACGACAGTCCGGGCTGAGTCCGTCGACTACGACAGGCTGCCG 503985
Qy 166 AlaArgIleYrYsIleValAlaYrAsnGlnProYrIleAsnYrSerAlaGln 185
Db 503986 GTGTCCGGTCCGGGTGGGTGAGTATCGCGGCGCTGTCTCTTTTGGC----- 504033
Qy 186 TyrArgThrGlnGlyAlaValAlaGluAlaIleValAlaGluAlaSerLeuIleArg 205
Db 504034 -----CAGAAAGAACAGCAGCGCGCGAGCGGCTCGGTCGATCATTTGCT 504084
Qy 206 -----SerValAlaSerPheSerIleYrSerProHisIleGlnGlu 221
Db 504085 GACAACATCGACGACGAGCGGATGGCGGACCTCGGGGCTATATCCAGCTC----- 504138
Qy 222 TyrGlnAspGlyValProYrIleProThrAlaIleYrValGluAspAlaGluMet 241
Db 504139 -----AAGATCCCGGTGAGTACACCAAG----- 504165
Qy 242 MetSerArgMetAlaSerHisglYIleYrIleValIleGlnLeuYs----- 257
Db 504166 -----TCGGTCGATTCAGCTACCGGACAGTCTGGG 504198
Qy 258 ---MetGlyAlaYrThrYrProAspThrAspSerPhe-----AsnThrValAla 273
Db 504199 CCAACACACCGTCAGAGCTACGCGGACGACCAAGTTTCAAGCGCGCAACGTCATCGCG 504258
Qy 274 GluIle---ThrGlySerYrYrProGluGlnValValLeuValSerGlyHisLeuAsp 292
Db 504259 CAGACGAAGCGGGGTGTGTG-----GCCAACCTGTGTATGAGGAGTGGCATTTTGGAC 504312
Qy 293 SerThrAspValGlyGlnGlyAlaMetAspAspIleGlyGlyAlaPheIleSerTrpGlu 312
Db 504313 AGCGTTCGGAAGGACCGCGCATACAGACACCGCTCGGAGTGGCT-----GCG 504363
Qy 313 AlaLeuSerLeuIleYrAspLeuGlyLeuArgPro-----IysArgThrLeuArgLeu 330
Db 504364 GTTCTGAAACGCGCAGTGCAGCTGGGAACTCACCGCATGTGTCAACGCGGTACGGTTC 504423
Qy 331 ValLeuThrPheAlaGluGlnGlnGlyValGlyAlaPheGlnIleYrYrGlnLeuHis 350
Db 504424 GCCTTCGGGGCGCGGAAATTCGGCTGATGGCTACAGAACTACGTGATCGCTG 504483
Qy 351 LysValAsn-----IleSerAsnYrSerLeuValMetGluSer 363
Db 504484 GACATCGACGCGCTCAAAAGCATCGCGCTGATGTGAATTCGACATGTTGGCGTCCGCG 504543
Qy 364 AspAlaGlyThrPheLeuProThrGly-----LeuGlnPheThrGlySer 378
Db 504544 AACCGGGTTACTTCACTACGACGAGTACCAAGTCCGCTGCGCTAGAACGCGCGGTACG 504603
Qy 379 -----GluYrAlaArgAlaIleMetGluGluValMetSerLeuLeuGlnPro 394
Db 504604 CCGGTGTGCCCCGAAGGCTCGCGGTATCGAGCGCAAGTTCGCTGCTATTAAGATG 504663
Qy 395 Leuamnllethr---GlnValLeuSerHisglYleuGluYThrAspIleAsnPhetripIle 413
Db 504664 GCCGGCAAGACCGCGGAGACACCTCGTTCGACGGTGGTCCGACTACGAGCGCTTCACG 504723
Qy 414 GlnAlaGlyValProGlyAlaSerLeu----- 422
Db 504724 CTGGCGGGTATCCCTTCGGGTGGCTGTTCCTCGGCGGTAGTCAGAAAGTCCGCCGAG 504783
Qy 423 -----LeuAspAspLeuYrYrYrPhe 430
Db 504784 CAAGCCAGCTCTGGGGCGGACCGCGACGAGCTTCGATCCCAACTAT----- 504834
Qy 431 PheHeHisSerHisglYAspThrMetThrValMetAspProlysglMet----- 448
Db 504835 -----CACCAAGAACAGACACCTCGACCATATCGACCGGACCGCGCTGATTC 504885
Qy 449 AsnValAlaAlaValAlaValAlaValAlaValSerYr 460

```

[illegible][illegible]

Db	1902	TATTCTGGAAATCCAGAGATTCTTCTGTTTGGAGAGACAGATTATCTTATTG	1961
Qy	431	PhePheHisHisSerHis	436
		:::	
Db	1962	GG-TACCACCATGGACAC	1978

Search completed: December 23, 2003, 02:25:48  
Job time : 2472 secs

**THIS PAGE BLANK (USPTO)**



MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/745,763  
FILING DATE: 18-Jun-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Springer, Suzanne A.  
REGISTRATION NUMBER: 41,323  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8284  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1851 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 35:  
US-09-745-763-35  
Alignment Scores:  
Pred. No.: 3,756-286 Length: 1851  
Score: 2417.00 Matches: 472  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
US-09-745-763-36 (1-472) x US-09-745-763-35 (1-1851)  
QY 1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGly 20  
DB 99 ATGAATTCCTTATCTTCGATTTTTCGGTGGTTCACCTTTTATCCCTGCTGCTGG 158  
QY 21 LysAlaIleCysValAsnGlyIleSerLysArgThrPheGluGluIleValGlyGlu 40  
DB 159 AAGCTATATGCAAGATGCACTCTTAAGAGACTTTTGAAGAAATTAAGAAATTA 218  
QY 41 AlaSerCysGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60  
DB 219 GCCACGTGGAGATGTTGCTTAAGCAATCATCAACCTTATGTTAAAGCCCG 278  
QY 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
DB 279 AACGATCTTATGACGATTTGGCAGCTTCTGTGTGATCTGTGGACCCAGACTGAGTGC 338  
QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100  
DB 339 TCCAGAACCTTAGAAGAACCCATCCAAATTATGACCAAAACCTGAGCAAGATGGCTG 398  
QY 101 GluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluGluSerAla 120  
DB 399 GAGAAAGTTCACTGAGCCAGTGAAGATACCCCACTGGGAGAGGAGAAAGATCAGCT 458  
QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLysSerIleGly 140  
DB 459 GTGATGCTGAGCCCAAGATTTCAATGAATAGCATCTGGGTCTTGGCAGCAGCTTGG 518  
QY 141 ThrProProGluGlyIleThrAlaGluValLeuValIleThrSerPheAspGluLeuGln 160  
DB 519 ACTCTCCAGAAAGCAATTACAGCAGAGTCTGGGTGAGCACTCTTCATCAATGCGAG 578  
QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValIleTyrAsnGlnProTyrIleAsnTyr 180  
DB 579 AGAAGGCTCAGAGCAAGAGAGAGATGTTGTTTAAACCAACCTTACATCACTAC 638  
QY 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaIleLysValGlyAlaLeu 200  
DB 639 TCAAGGAGCGTGCATATACGAAACGAGGAGGCGGTGAGAGCTGCCAAGGTGGGCTTGG 698

QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220  
DB 699 GCATCTCTCATTCATCCGTGGCTCTCTTCCATCTACAGCTCTCACACAGGATTCAG 758  
QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
DB 759 GAATACCGAGATGGCGTGGCCCAAAATTCAAACACCTGTATTAACGTGGAGAGCAGAA 818  
QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
DB 819 ATGATGTCAGAAATGGCTTCTCATGGAGATCAAAATGTTCATTCAGCTAAAGATGGGCA 878  
QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280  
DB 879 AAGACCTACCAAGTATCATGATCTTCTTCAACCTTACAGAGATCATCTGGAGCAATAT 938  
QY 281 ProGluGlnValIleValLeuValSerGlyHisLeuAspSerTyrAspValGlyGlnGlyAla 300  
DB 939 CCAGAACAGTGTGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 998  
QY 301 MetAspAspGlyGlyGlyAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320  
DB 999 ATGATGATGGCGGTGAGCTTTATATCATGGAGACACTTCACTTATTAAGATCTT 1058  
QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTyrPThrAlaGluGluGlnGly 340  
DB 1059 GGGCTGCTCCAAAGAGAGACTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1118  
QY 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360  
DB 1119 GTTGCTGCTTCCAGTATATCATGATGATGATGATGATGATGATGATGATGATGATGATG 1178  
QY 361 MetGluSerAspAlaGlyTyrPheLeuProThrGlyLeuGlnIleThrGlySerGlyLys 380  
DB 1179 ATGAGATCTGACGAGAGAACTTCTTACCCACTGGCTGCAATTCACCTGAGGAGAAAG 1238  
QY 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400  
DB 1239 GCCAGGCCATCATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1298  
QY 401 LeuSerHisGlyGluGlyThrAspIleAsnThrPheIleGlnAlaGlyValProGlyAla 420  
DB 1299 CTGAGCCATGAGAGAGGACGACATCATCACTTTGGATCCAAAGCTGGAGTCCCTGGAGCC 1358  
QY 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440  
DB 1359 AGTCTACTGTGACTTATATCAAGATTTCTTCTTCAATCACTCCACGAGACCAATG 1418  
QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaValIleThrAlaValIleSerTyr 460  
DB 1419 ACTGTCTGATCTCCAAAGCAGATGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1478  
QY 461 ValValAlaAspMetGluMetLeuProArgSer 472  
DB 1479 GTTGTCAGACATGGAAGAAATGCTGCTGATGCTC 1514  
RESUR 2  
US-09-984-271-28  
; Sequence 28, Application US/09984271  
; Publication No. US20030040088A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 71 Human Secreted Proteins  
; FILE REFERENCE: P2030P1  
; CURRENT APPLICATION NUMBER: US/09/984, 271  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 09/482, 273  
; PRIOR FILING DATE: 2000-01-13  
; PRIOR APPLICATION NUMBER: PCT/US99/15849  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: 60/092, 921  
; PRIOR FILING DATE: 1998-07-15

PRIOR APPLICATION NUMBER: 60/092,922  
PRIOR FILING DATE: 1998-07-15  
PRIOR APPLICATION NUMBER: 60/092,956  
PRIOR FILING DATE: 1998-07-15  
NUMBER OF SEQ ID NOS: 267  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 28  
LENGTH: 1863  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-984-271-28

Alignment Scores:  
Pred. No.: 3,79e-286 Length: 1863  
Score: 2417.00 Matches: 472  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 11 Gaps: 0

US-09-745-763-36 (1-472) x US-09-984-271-28 (1-1863)

QY 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCyseSerGly 20  
DB 99 ATGAAATTCCTTATCTTCGCACTTTTCGCTGCTTCACTTTATCCCTGCTGCGG 158  
QY 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40  
DB 159 AAACCTATATCAAGAAATGSCATCTCTAAGAGACTTTTGAAGAAATTAAGAAATA 218  
QY 41 AlaserCyseGlyAspValAlaLysAlaIleIleAsnLeuAlaValTYRGLYValaGln 60  
DB 219 GCCAGCTGTGAGAGATGTTGCTAAGCAATCATCACTAGCTGTTTATGTAAGCCAG 278  
QY 61 AsnArgSerTYRGLYValGluLeuAlaLeuValAspThrValGlyProArgLeuSerGly 80  
DB 279 AACGATCTCATAGCGATGCGATGCGACTCTGCTGTTGACTGTTGAGCCAGACTAGTGGC 338  
QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTYRGLYAsnLeuGlnAspGlyLeu 100  
DB 339 TCCAGAACTCAGAAAGAAAGCAATCCAAATTAATGTAACAAACCTGCGAGCAAGATGGGCTG 398  
QY 101 GluLysValHisLeuGluProValArgIleProHisTrpGluArgGlyGluGluSerAla 120  
DB 399 GAGAAAGTTCACTGAGAGCCAGTGAATACCCCACTGAGAGAGGAGAAAGATCACT 458  
QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
DB 459 GTGATGCTGAGCAAGAAATTCATTAAGATACCATCTCGGCTTGGCAGCAGCATTTGGG 518  
QY 141 ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160  
DB 519 ACTCTCTCAGAAAGGATTACAGCAAGAACTTCTGGGTGGAGCTCTTTCAGAACTGAG 578  
QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTYRAsnGlnProTYRILEAsnTYR 180  
DB 579 AGAAGGCTCAGAAAGCAAGAGGAAGTTGTTGTTTAACCAACTTAATCATCACTAC 638  
QY 181 SerArgThrValGlnTYRArgThrGlnGlyAlaValGluAlaIleLysValaGlyAlaLeu 200  
DB 639 TCAAGAGCGGTGCAATACCAAGCGAGGGGGGTGGAAGCTGCCAAGTTGGGGCTTTG 698  
QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTYRSerProHisThrGlyIleGln 220  
DB 699 GCATCTTCATTCGATCCGTGGCTCCCTTCTCCATCTCACTCACTCACTCACTCACTCA 758  
QY 221 GluTYRGLYAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
DB 759 GAATTCAGAGATGGCTGCCCAAGTTCCAAAGCTTATTAACGTTGGAAGATGCAGAA 818  
QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
DB 819 ATGATGTCAAGAAATGCTTCTCATGTGATCAAAATTTTCATTCAAGCTTAAGATGGGGCA 878

QY 261 LysThrTYRProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTYR 280  
DB 879 AAGACCTACCCAGATCTGATGTTCTTCAACACTGTAGCAGATCACTGGAGCAAAATAT 938  
QY 281 ProGluGlnValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300  
DB 939 CCAGAACAGGTGTACTGTCAGTGCACATCTGCAGCTGGAGATGTTGGGAGGGTGGC 998  
QY 301 MetAspAspGlyGlyValaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320  
DB 999 ATGAGATGCGCGGTGAGCTTTATATCATGGAAGCACTCTCATTTAAGAATCTT 1058  
QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTrpThrAlaGluGlnGlyGly 340  
DB 1059 GGGCTGTCTCAAGAGACTCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1118  
QY 341 ValGlyAlaPheGlnTYRGLYValHisLysValaAsnIleSerAsnTYRSerLeuVal 360  
DB 1119 GTTGTGCTCTTCCAGTATATATCAAGTTACACAGGTAAATTTCCAACTACAGCTGCTG 1178  
QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380  
DB 1179 ATGAGCTGACCGCAGAACCTTCTTACCACTGGGCTGCAATTCACCTGCGAGTAAAG 1238  
QY 381 AlaArgAlaIleMetGluGluValaMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400  
DB 1239 GCCAGGCGCATCATGAGAGGTTATGAGCTGCTGCAAGCCCTCAATATCACTCAAGTTC 1298  
QY 401 LeuSerHisGlyGluGlyTYRAspIleAsnPheTrpIleGlnAlaGlyValaProGlyAla 420  
DB 1299 CTGAGCCATGAGAGAGGACAGACATCACTTTTGATCCAACTGAGAGCTGGAGCC 1358  
QY 421 SerLeuLeuAspLeuTYRGLYTYRPhenPheHisHisSerHisGlyAspThrMet 440  
DB 1359 AGTCTACTGATGATCATTAACAATATTTCTTCCATCACTCCCAAGAGACACCATG 1418  
QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaValaTrpAlaValaLysTYR 460  
DB 1419 ACTGTCACTGATCCAAAGCAAGATGATGTTGCTGCTGCTGCTGCTGCTGCTGCTTAT 1478  
QY 461 ValValaAlaAspMetGluGluMetLeuProArgSer 472  
DB 1479 GTTGTTCACACATGAGAAAGATGCTGCTAGTCC 1514

RESULT 3  
US-09-917-800A-505  
Sequence 505, Application US/09917800A  
Patent No. US20020119462A1  
GENERAL INFORMATION:  
APPLICANT: Mendrick, Donna  
APPLICANT: Porter, Mark  
APPLICANT: Johnson, Kory  
APPLICANT: Caesle, Arthur  
APPLICANT: Blaschoff, Michael  
APPLICANT: Gene Logic, Inc.  
TITLE OR INVENTION: Molecular Toxicology Modeling  
FILE REFERENCE: 44921-5038-US  
CURRENT APPLICATION NUMBER: US/09/917,800A  
PRIOR FILING DATE: 2001-07-31  
PRIOR APPLICATION NUMBER: US 60/222,040  
PRIOR FILING DATE: 2000-07-31  
PRIOR APPLICATION NUMBER: US 60/222,880  
PRIOR FILING DATE: 2000-11-02  
PRIOR APPLICATION NUMBER: US 60/230,029  
PRIOR FILING DATE: 2001-05-11  
PRIOR APPLICATION NUMBER: US 60/230,645  
PRIOR FILING DATE: 2001-05-15  
PRIOR APPLICATION NUMBER: US 60/232,336  
PRIOR FILING DATE: 2001-05-22  
PRIOR APPLICATION NUMBER: US 60/235,798  
PRIOR FILING DATE: 2001-06-06  
PRIOR APPLICATION NUMBER: US 60/297,457

; PRIOR FILING DATE: 2001-06-13  
 ; PRIOR APPLICATION NUMBER: US 60/298,884  
 ; PRIOR FILING DATE: 2001-06-19  
 ; PRIOR APPLICATION NUMBER: US 60/303,459  
 ; PRIOR FILING DATE: 2001-07-09  
 ; NUMBER OF SEQ ID NOS: 1740  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 505  
 ; LENGTH: 1778  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AF097723  
 US-09-917-800A-505

Alignment Scores:  
 Pred. No.: 3,316-249 Length: 1778  
 Score: 2116.00 Matches: 413  
 Percent Similarity: 92.58% Conservative: 24  
 Best Local Similarity: 87.50% Mismatches: 35  
 Query Match: 87.55% Indels: 0  
 DB: 10 Gaps: 0

US-09-745-763-36 (1-472) x US-09-917-800A-505 (1-1778)

QY 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCySerGly 20  
 DB 121 ATGAGGCTCTCTTTCTTCTCTGTCGTGCTGTTCACTTTCTCTGCGCTCTGGA 180  
 QY 21 LysAlaIleCyLeuAsnGlyIleSerLysArgThrPheGluGluIleLysGluIle 40  
 DB 181 AAAGCTATATACAAAGAGTGTCTTCTCAGCAATTTCAAGAAATATAAAGAAATA 240  
 QY 41 AlaSerCySerGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyValAlaGln 60  
 DB 241 GCCAATGATGAGATGTTGCTTAAGCAATATCACTTGTGTTATGAAATATACAG 300  
 QY 61 AsnArgSerTyrGlyArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
 DB 301 AACCGGTGATGAGCGCTTGGGACTTCTAGTGTACTGTTGGACCCAGACTGAGTGGC 360  
 QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100  
 DB 361 TCTAAGACCTATGAGAAAGCTATCCAAATCATGTATCCAAACCTGCAACAGATGGCTG 420  
 QY 101 GluLysValHisLeuGluProValArgIleProHisTTPGluArgGlyGluGluSerAla 120  
 DB 421 GAAAACGTCCACCTGAGCAGGTCAGAAATCACTCACTGGGGAGGGGGAAGAAATCTCA 480  
 QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
 DB 481 GTGATGGTGGTGGCTCGAATTCACAAAGTTGGCTATTTTAGGCTTGGCGGACGCTTGG 540  
 QY 141 ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160  
 DB 541 ACTCTCTCTGAAGATATCACAGCAAGATACGTGGGGCCCTTTTGTGAATCTCA 600  
 QY 161 ArgAlaGluAspGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180  
 DB 601 AGAAGGCAATCAGAGCAGAGGAAAGATGTGTTGTATTAACACCTTACACCTGACAT 660  
 QY 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaLysValGlyValLeu 200  
 DB 661 GGGAAATCTGTCAATACCGGAGCGGAGCTGTGGAAAGCTGCCAAGGTGGGGCCGTG 720  
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220  
 DB 721 GCATCCCTCATCCAGTACAGTACTTTTTCATCTACAGTCCCTCACACAGGTATCA 780  
 QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCySerIleThrValGluAspAlaGlu 240  
 DB 781 GGATATCAAGATGTGTGCTCCCAAGATTCACACAGCTGTATCAATAGAAATGACAGAA 840

QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
 DB 841 ATGATGCTCGAATGGCTTCCGTGGGACAAATTTCAATTCATTCGAAATGGAAGCA 900  
 QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280  
 DB 901 AAGACCTATCCAGATACAGATTCCTTCACACTTTCAGAGATCACTGGAGCAAGATAT 960  
 QY 281 ProGluIleValIleValIleValIleSerGlyHisLeuAspSerThrAspValGlyGlnGlyAla 300  
 DB 961 CCAGAGAAATGTTCTCTGCTCAGTGCATCTGCACAGCTGGACGCTGGGAGGTCGA 1020  
 QY 301 MetAspAspGlyGlyGlyAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320  
 DB 1021 CTGATGATGCGCGGTGAGGCTTCAATCATGGGAACACCTCATCTTGTAAAGATCTT 1080  
 QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrAlaGluGluGlnGlyGly 340  
 DB 1081 GGGCTGCGTCCAAAGAGACTCTGCGCTGTGTCTGCAACCGCAAGAACAGAGAGG 1140  
 QY 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360  
 DB 1141 GTTGTGCTCCCAATATATGAGCTACATTAAGCAAAATTTTCCAAATACAGTTTGGTG 1200  
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380  
 DB 1201 ATGAGGCTGATCTGAGAACCTTCTTACCACTGGGTGAGTTCACCGGAGATGACAG 1260  
 QY 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400  
 DB 1261 GCCGAGGCTATCAAGAAAGATCATAGTCTCTGCAACCCCTCAATATCACCAAGATC 1320  
 QY 401 LeuSerHisGlyGluGlyIleThrAspIleAsnPheThrIleGlnAlaGlyValProGlyAla 420  
 DB 1321 TTTATATATCAGAGAAAGTCACTGACATTAATCTTGGATTCAGCTGAGTGGAGCC 1380  
 QY 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisIleSerHisGlyAspThrMet 440  
 DB 1381 AGTCTGGAGATGATCTTGTACAAGTATTTCTTTTTCATCATTCCTCAGTGGAGACCATG 1440  
 QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaValThrAlaValSerTyr 460  
 DB 1441 ACTGCCATGATCCAAAGAGATGAATGTGCTGCTCTCTTGGCTGTGTGCTTAC 1500  
 QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472  
 DB 1501 GTTGTGCAAGCATGAGAGAAATGCTGCCAGGTCC 1536

RESULT 4  
 US-09-984-271-95  
 ; Sequence 95, Application US/09984271  
 ; Publication No. US20030040088A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: 71 Human Secreted Proteins  
 ; FILE REFERENCE: P2030P1  
 ; CURRENT APPLICATION NUMBER: US/09/984,271  
 ; CURRENT FILING DATE: 2001-10-29  
 ; PRIOR APPLICATION NUMBER: 09/482,273  
 ; PRIOR FILING DATE: 2000-01-13  
 ; PRIOR APPLICATION NUMBER: PCT/US99/15849  
 ; PRIOR FILING DATE: 1999-07-14  
 ; PRIOR APPLICATION NUMBER: 60/092,921  
 ; PRIOR FILING DATE: 1998-07-15  
 ; PRIOR APPLICATION NUMBER: 60/092,922  
 ; PRIOR FILING DATE: 1998-07-15  
 ; PRIOR APPLICATION NUMBER: 60/092,956  
 ; PRIOR FILING DATE: 1998-07-15  
 ; NUMBER OF SEQ ID NOS: 267  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 95  
 ; LENGTH: 1134  
 ; TYPE: DNA



ORGANISM: Homo sapiens  
US-09-984-271-95

Alignment Scores:  
Pred. No.: 5,99e-154 Length: 1134  
Score: 1338.00 Matches: 261  
Percent Similarity: 99.62% Conservative: 0  
Best Local Similarity: 99.62% Mismatches: 1  
Query Match: 55.36% Indels: 1  
DB: 11 Gaps: 0

US-09-745-763-36 (1-472) x US-09-984-271-95 (1-1134)

```

Qy 211 SerIleTyrSerProH1ethrGlyIleGlnGlnIleuTyrGlnAAspGlyValProIleS1lepro 230
Db 1 TCcATCTACAGCTCCTCACAGAGTATTCAGAAATACCGAGATGGCGGCTCCCAAGATTCa 60
Qy 231 ThrAlaCysIleThrValGluAspAlaGluMetSerArgMetAlaSerHisGlyIle 250
Db 61 ACAGCCGTGATTTACGTCGAGAGATGACAGAAATGATGTCAGAAATGGCTTCATGGGATC 120
Qy 251 LysIleValIleGlnIleuLysMetGlyValAlaYsThrTyrProAspThrAspSerPheAsn 270
Db 121 AAAATTGTCATTCAGCTAAAGATGGGGGCAAAAGCTTACCAGATCTGATTCCTTCAAC 180
Qy 271 ThrValAlaGluIleThrGlySerIleTyrProGluGlnValValIleuValSerGlyHis 290
Db 181 ACTGTACAGAGATCATCTGGAGCAAAATATCCAGAACGGTTGACTGCTGATGAGACAT 240
Qy 291 LeuAspSerTyrAspValGlyGlnGlyValMetAspAspGlyGlyValAlaPheIleSer 310
Db 241 CTGACACAGCTGGAGTGGGCAAGGTCATGATGATGCGGAGGCTTTATATCA 300
Qy 311 TrpGluAlaLeuSerIleuIleLysAspLeuGlyLeuAspGlyPheGlyThrLeuArgLeu 330
Db 301 TGGGAACACCTCTCACTTAAAGATCTTGGCGCTCCCAAGAGAGACTCGCGGTG 360
Qy 331 ValLeuTrpThrAlaGluGlnGlnGlyValGlyValAlaPheGlnIleTyrTrpGlnLeuHis 350
Db 361 GTGCTCTGAGCTGCAGAAACAGAGTGGAGTTGGCTTCCAGTATTTACATTAC 420
Qy 351 LysValAsnIleSerAsnTyrSerLeuValMetGluSerAspAlaGlyThrPheLeuPro 370
Db 421 AAGTTAATATTTCCACTACAGTCTGATGATGAGTGTGACGAGAACTTCTTACC 480
Qy 371 ThrGlyLeuGlnPheThrGlySerGlyLysAlaArgAlaIleMetGluGlnValMetSer 390
Db 481 ACTGGGCTGCATTCACCTGCGAGTCAAAAGCCAGGCGC-ATCATGAGAGAGTTATGAGC 539
Qy 391 LeuLeuGlnProLeuAsnIleThrGlnValLeuSerHisGlyGlyGlnIleTyrThrAspIleAsn 410
Db 540 CTGCTGAGCCCTCAATATCATCTCAGTCCCTGAGCCATGGAAGGAGACACATCAAC 599
Qy 411 PheTrpIleGlnAlaGlyValProGlyValaSerLeuLeuAspAspLeuTyrLysTyrPhe 430
Db 600 TTTTGGATCCAGCTGAGTCCCTGAGCCAGTCTACTTACTTATGACTTATACAGATTTCC 659
Qy 431 PhePheHisHisSerHisGlyAspThrMetThrValMetAspProLysGlnMetAsnVal 450
Db 660 TTCTTCCATCATCTCCAGGAGACCCATGATGATGATGATGATGATGATGATGATGAT 719
Qy 451 AlaAlaAlaValAlaValAlaValSerTyrValAlaAlaAspMetGluGlnMetLeuPro 470
Db 720 GCTGTGCTGTGTGGCTGTGTGTGTCTTATGTGTGTGACAGATGAGAAATGCTGCT 779
Qy 471 ArgSer 472
Db 780 AGGTCC 785

```

RESULT 5  
US-09-833-381-1929  
; Sequence 1929, Application US/09833381  
; Patent No. US20020132090A1

GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1929  
; LENGTH: 357  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(357)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1929

Alignment Scores:  
Pred. No.: 7.58e-53 Length: 357  
Score: 508.50 Matches: 98  
Percent Similarity: 92.37% Conservative: 11  
Best Local Similarity: 83.05% Mismatches: 7  
Query Match: 21.04% Indels: 2  
DB: 10 Gaps: 1

US-09-745-763-36 (1-472) x US-09-833-381-1929 (1-357)

```

Qy 321 GlyLeuArgProIlyAspArgThrLeuArgLeuValLeuTrpThrAlaGluGlnGlnGly 340
Db 7 GGCCTCCGA---AAGAGAGCTCTGCGCTGCTCTGAGCTGCAGAAAGAGAGAGA 63
Qy 341 ValGlyAlaPheGlnIleTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360
Db 64 ATGTGCTCTCCCGATTTATGAGCTCATTAAGCAAAATATTTCCAGTACAGCTGGTG 123
Qy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380
Db 124 ATGGAAGCTGACTCAGAGAACCTTCTTACCACTGAGCTGCAAGTTCGACAGTGAAG 183
Qy 381 AlaArgAlaIleMetGluGlnValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db 184 GCCAGGCTATCATGAAGAGAGATCATATCTTCTGCAACCCCTCAATGTCACCAAGGTC 243
Qy 401 LeuSerHisGlyGlnIleTyrAspIleAsnPheTrpIleGlnAla-GlyValProGlyAl 420
Db 244 TTATGTAATGAGAGAGAGATGATATTAATCTTCTGATCCAGCTTGAGATNCTGGAGC 303
Qy 420 AserLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGly 437
Db 304 CAGCTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 355

```

RESULT 6  
US-09-833-381-1930  
; Sequence 1930, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1930  
; LENGTH: 427  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:

NAME/KEY: misc feature  
LOCATION: (1)-(427)  
OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1930

Alignment Scores:  
Pred. No.: 1,12e-44 Length: 427  
Score: 443.00 Matches: 91  
Percent Similarity: 85.84% Conservative: 6  
Best Local Similarity: 80.53% Mismatches: 14  
Query Match: 18.33% Indels: 2  
DB: 10 Gaps: 1

US-09-745-763-36 (1-472) x US-09-833-381-1930 (1-427)

Qy 1 MetlysrpHeuLlePheAlaPheheGlyValHIsleuLeuSerLeuCySerGly 20  
Db 91 ATGAGGTCTCTTTCTTCTCTGTTCAAT-----GTTCACTTTTAGCCTTGGCTGTGT 144  
Qy 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleGlySerGly 40  
Db 145 AAAGCTGATTCAAAATGCTTTCTCTCGCAACATTTCGAAATTAAGAAAGAAATTA 204  
Qy 41 AlaSerCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60  
Db 205 GCCAATGATGAAGATGTTGCTTAAGCAATTATCAACCTTGTATTATGTAATATACAG 264  
Qy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
Db 265 AACCGCTCTATAGACGCTTGGACTTCTAGTTGATGACTGTGAGCCAGACTGAGTGGC 324  
Qy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnGlnAspGlyLeu 100  
Db 325 TCTAAGAACTTGAAGAACTTATCAATATCATATGTTACCAAAACCTTGACAGATGGCTTA 384  
Qy 101 GluLysValHisLeuGluProValArgIleProHisTyr 113  
Db 385 GAAATGTTCACCTGAGAGAGTCAANAATACCCCACTGG 423

# RESULT 7

US-09-974-300-2337  
Sequence 2337, Application US/09974300  
Patent No. US20020146721A1  
GENERAL INFORMATION:  
APPLICANT: Berka, Randy M.  
TITLE OF INVENTION: Methods for Monitoring Multiple Gene  
FILE REFERENCE: 10085,500-US  
CURRENT APPLICATION NUMBER: US/09/974,300  
PRIOR FILING DATE: 2001-10-05  
PRIOR APPLICATION NUMBER: 09/680,598  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: 60/279,526  
NUMBER OF SEQ ID NOS: 8481  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 2337  
LENGTH: 1346  
TYPE: DNA  
ORGANISM: Bacillus licheniformis  
US-09-974-300-2337

## Alignment Scores:

Pred. No.: 8.09e-13 Length: 1346  
Score: 190.00 Matches: 110  
Percent Similarity: 40.28% Conservative: 89  
Best Local Similarity: 23.27% Mismatches: 215  
Query Match: 7.86% Indels: 80  
DB: 10 Gaps: 23

US-09-745-763-36 (1-472) x US-09-974-300-2337 (1-1346)

Qy 2 LysPheLeuIlePheAlaPheheGlyValHIsleuLeuSerLeuCySerGlyLys 21  
Db 10 AAATGATGATGTTCCGATTC-----GGCTATGATCATTTGCGGGCC 54  
Qy 22 AlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleGlySerGly 41  
Db 55 GTGCTCCCGCATGGA-----ACGGGAATGCACCTCAGCGGCCCTCAGAAACAGCC 108  
Qy 42 SerCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAla----- 59  
Db 109 ATCCGAAAAGATATCGAAAA-----TTACGCAAAAAAATTCAC 147  
Qy 60 GlnAsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSer 79  
Db 148 GAGAACCGCGCTTATCAAGCATTTTACCATTTTAAAGCAAAAGATGGCGCGGTGCA 207  
Qy 80 GlySerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGly 99  
Db 208 GGCACGCGCAGAAAGAAAAAGAGCGCGCTTTATCCCTCAGATGAAATCAAT 267  
Qy 100 LeuGluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluSer 119  
Db 268 CTG--AAAGTACACACAAACTTACGATCTTACCGCGGTGAAAGG----- 315  
Qy 120 AlaValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIle 139  
Db 316 -----ACGCTTACCTTACGGAATATAGCTGCTCGCGCGCGCGCTTCCGCC 369  
Qy 140 GlyThrProProGluGlyIleThrAlaGluValLeu-----ValValThr 154  
Db 370 CCGACTCGACGAGAGGCTTGGCGCTCTCTATATGATGCGCGCTCGCGCTCGCGC 429  
Qy 155 SerPheAspGluLeuGlnArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsn 174  
Db 430 GACTTCACAGAA-----GAACGAAAGGAAAAATGCGCCGCT----- 465  
Qy 175 GlnProTyrIleAsnTyrSerArgThrValGlnTyrArgThrGlnGlyAlaValGluAla 194  
Db 466 -----ATTTTGAGAGAGAGAGCTGATCTTATGAAAAGCCAAAACGCTGAGCGCA 519  
Qy 195 AlaLysValGlyAlaLeuAlaSerLeuIleArgSerValAlaSerPheSerIleTyrSer 214  
Db 520 GCGCCAGCGGAGTATC-----ATTATATATACGTCGACACCTCGCTCGCTGACT 573  
Qy 215 ProHisThrGlyIleGlnGluTyrGlnAspGlyValProLysIleProThrAlaCysIle 234  
Db 574 CCGATCTTACCGGTAATTAAGTCGAT-----ATTCAGATTGCGCGCTC 618  
Qy 235 ThrValGlnAspAlaGluMetMetSerArgMetAlaSerHisGlyIleLysIleValIle 254  
Db 619 AAAAAGAGAGCGGGGAA-----AAGCTGCTTCTGAAACAAGAAGCATC----- 663  
Qy 255 GlnLeuLysMetGlyAlaLysThrTyrProAspThrAspSerPheAsnThrValAlaGlu 274  
Db 664 ---CTGAAGCTGAAGGCTCATTAATAAATCAACCTCGCAAAACGTAAATCGCGTCGCAAA 720  
Qy 275 IleThrGlySerLysTyrProGluGlnValValLeuValSerGlyHisLeuAspSerTyr 294  
Db 721 GCAAAAGGTGTCAAAATCTCGAC--ATGCTTATATGATGACTTCGCAATATGACAGCTT 777  
Qy 295 AspValGlyGlnGlyAlaMetAspAspGlyGlyGlyAlaPheIleSerTyrGluAlaLeu 314  
Db 778 CCATATGACCGGAGCCCAATGACAAATGCTCCGCGCACTTCAAGTCGCTTGAATCGGCC 837  
Qy 315 SerLeuIleLysAspLeuGlyLeuArgProLysArgThrLeuArgLeuValLeuTyrThr 334  
Db 838 CGAATTTTGAAGACA-----GTCCAGCCGCAACAAGAAATTCGCTTATATACATTCGCA 891  
Qy 335 AlaGluGluGlnGlyValGlyAlaPheGlnTyrTyr-----GlnLeu 349  
Db 892 GCCGAAAGAAATCGAGCTCTCGGTTCCGCAATATATGTCAGACATTTGCAAAATCGCAA 951  
Qy 350 HisLysValAsnIleSerAsnTyrSerLeuValMet-----GluSerAsp 364

```

Db      952 GTCAACGAGCGCTGGCACTTAATGATATGCGCAGCAAGCTGGAGAAATGCT 1011
Qy      365 ALaelyhrPheleupProthr-----GlyLeuGlnPheThylSeryl 379
Db      1012 TCACAGCTGATCAATATACCGCGAGCGTTGGCAACCTGCTGCTGACCTAAGTAA 1071
Qy      380 LysAlaArg--AlaIleMetGluValMetSerLeuGlnProLeuAsnIleThr 398
Db      1072 GCGGCTCTTAAACCTTGAAGAAAGACCTATATTTTACATCAAGCGGATCATCCGAC 1111
Qy      399 GlnValLeuSerHisGlyGlyThrAsp--IleAsnPhe--TrpIleGlnAlaGly 416
Db      1132 CATGTCCCATTCATGAGACCGCGATCGACTCGGCCAATTTATTTGGAGAGAG----- 1185
Qy      417 ValProGlyAlaSerLeuLeuAspAspLeuTyrlSerylPhePhePheHisSerHis 436
Db      1186 ---CCGGGACAGCGCGCTTTGGAGCTTGATC-----CACACCCCT 1224
Qy      437 GlyAspThrMetThrValMetAspProlyGlnMetAsnValAlaAlaValTrpAla 456
Db      1225 TACACACGATTAACATATGACCAAGACAGCGCTGAAGAACCGGCTCAATATCGCGGA 1284
Qy      457 ValValSerTyrlValAlaAlaAspMetGlnGluMetLeuPro 470
Db      1285 ACAGCTGTGTATTAATTCACCAAGAGGAAACAGAAACCG 1326

```

## RESULT 8

```

US-10-325-430-14
; Sequence 14, Application US/10325430
; Publication No. US2003015325A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; FILE REFERENCE: MP101-294P1RPM
; CURRENT APPLICATION NUMBER: US/10/325,430
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/341,953
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 2223
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2223)
US-10-325-430-14

```

## Alignment Scores:

```

Pred. No.:      3,13e-12      Length:      2223
Score:          188.00      Matches:      112
Percent Similarity: 35.24%      Conservative: 85
Best Local Similarity: 20.04%      Mismatches: 186
Query Match:      7.78%      Indels:      176
DB:              13      Gaps:      27

```

US-09-745-763-36 (1-472) x US-10-325-430-14 (1-2223)

```

Qy      58 LysAlaGlnAsn-----ArgSerTyrlGlnValGluAlaLeuLeuValAsp 72
Db      166 AAAGCTGAAAACATCAATCATTTCTGCTCTTTTACAAAGCTT----- 210
Qy      73 ThrValGlyProAlaGlyLeuSerGlySerIleAsn-----LeuGlnValAlaIleGln 89
Db      211 -----CTCATCTTGCGAGAGACAGAAACAAATTTCTGCTTGCCAGAAATCCAA 261
Qy      90 IleMetTyrlGlnAsnLeuGlnIleAsnGlyLeuGlnValIleLeuGlnPro---Val 108

```

```

Db      262 ACCCAGTGAAGAAATTTGACATAGATTACCAAGATTGGTTCATATATGATGCTCTTA 321
Qy      109 ArgIleProHisTrpGluValGlyGlyGluSerAlaValMetLeuGlnProAlaIleHis 128
Db      322 TCTTACCCCAATAGAGCAAAATGCCAATATATATGATGTGATGAA-----CAT 372
Qy      129 LysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGlyGlyIleThrAla 148
Db      373 GAACCTAGATTTTCAAAAC--TCATACCTTGAACACACACAGATGCTATGAGAT 429
Qy      149 GluValLeuValThrSerPheAspGlyLeuGlnAlaArgAlaSerGlyAlaArgGly 168
Db      430 GTTACAAATATTTGGCCACCATATATATGCTTCTCAGCCCAAGCGATG---CCAGAGGA 486
Qy      169 LysIleValValTyrlAsnGlnProTyrlLeuAsnTyrlSerTyrlGlyThrValGlnTyrl 186
Db      487 GATCTTGTGTA-----TATGTAACTATGCTGCACTGCAAGACCTTTTCAAA 531
Qy      187 -----ArgThrGlnGlyAlaValAlaGluAlaIleValValGlyValAla 199
Db      532 CTGAAAGAGAGATGGGATCATCACTGATCTGGAAGATTTGATATGCAAGATATGAAAA 591
Qy      200 Leu-----AlaSerLeuIleArgSerValAlaSerPheSerIle 212
Db      592 ATCTTCAGAGGAATATAAATTAAATGATCCATGTATACAGAGCCATAGAAATCATCTTG 651
Qy      213 TyrlSer----- 214
Db      652 TACTCAGATCCAGCTGATCTTCTGCTGAGAGTACAGCCATATCCCAAGATGGAAT 711
Qy      215 ---ProHisThrGlyIleGln----- 220
Db      712 CTTCCTGGAAGCTGACGCCAGAGAGAAATGTGTTAAATTTGAATGCTGCTGATACCA 771
Qy      221 -----GluTyrlGlnAsp 224
Db      772 CTCACCTCCAGGCTATCCAGCAAAAGATATACCTTTCAGACTGTATGTTGAAGAGAGTG 831
Qy      225 GlyValProlySrlleProThrAlaCysIleThrValAlaGluAspAlaGluMetSerArg 244
Db      832 GGAATCCCCCGAATACCTGTATATATCCATCCATGATATATATATGAGAAATATATATTCGC 891
Qy      245 MetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAlaLysThr----- 262
Db      892 TACTTGGGA---GGAATGCTCCACAGATPAGAGTTGGAAGAGGCCCTTAATGTGAGT 948
Qy      263 -----TyrProAspThrAspSerPhe----- 269
Db      949 TATAGTATCGGACCTGGCTTTACAGGAGATGATTTCTTTCAGAAAGCTTAGAATGCATGTT 1008
Qy      270 -----AsnThrValAlaGluIleThrGlySerLys 279
Db      1009 TATTAACATCAATAAATTACAGAGATTATACATATGATGTAAGTATCAGAGGATCTGTG 1068
Qy      280 TyrProGluGlnValValLeuValSerGlyHisIleuAspSerTrpAspValGlyGlnGly 299
Db      1069 GAACCTGACGAGTATGTATATCTGAGAGTACCGGAGCTCTCTGGATATT-----GGA 1122
Qy      300 AlaMetAspAspGlyGlyGlyAlaPheIleSerTrpGlu-----AlaLeuSerLeu 316
Db      1123 GCTATTTGACCAACACAGATGGGCTGCTGTTTTCAGAAABAATTCGCCGAGATTGGAAAA 1182
Qy      317 IleLysAspLeuGlyLeuArgProlyArgThrLeuArgLeuValLeuTrpThrAlaGlu 336
Db      1183 CTGATGTAGTAAAGCTGAGAGCTTGAAGAAACTATATTTTTCACCTGAGATGACGAA 1242
Qy      337 GluGlnGlyGlyValGlyAlaPheGlnTyrlGlnIleuHisLysValAsnIleSerAsn 356
Db      1243 GAAATTTGACTTCTGGGTTCCAGAAATGGCTAG-----GAGAAAT 1284
Qy      357 TyrlSerLeuValMetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThr 376

```

```

Db      1285 GTCAAAATCTCCAGAGAGACAT---GCTATATCACTCGATTATCATAGAA 1341
Qy      377 GYserGluYsAlaArgAla-----lleMeGluGluValMetSerleu 391
Db      1342 GGCATTTACTCTCAGAGTTGACTACTCCCTCTTTACCAATTAGTATTAAGTGTG 1401
Qy      392 LeuGlnProLeu----- 395
Db      1402 ACAAAGAGATCCCAAGCCCTGATGATGGTTTGAGAGTAATCACTGTATGAAGCTGG 1461
Qy      396 -----AenllethrglnValleuSerHisGlyGlu 405
Db      1462 TTGAAAAAGACCTTCACTGAAATAATAATTGGCTAGATCAATGAAGCTGGAGTCT 1521
Qy      406 GYThrAspPheAsnPhetripIleGln---AlaGlyValProGlyAla----- 420
Db      1522 GGAAGTGTCTTGAAAGCTTATTTTCAAGACCTTGAAATGCTTCAAGCAGACCCGTTAC 1581
Qy      421 -----SerleuLeuAspAspLeuTyrIleTyrPhePheHisSerHisGlyAsp 438
Db      1582 ACTAAGATTAAGAAAAACAGATAGTACAGACGACTCCAGCTGTACACATTTATGAG 1641
Qy      439 ThrMetThrValMet-----AspPro-----LysGlnMetAsnValAla 451
Db      1642 ACATTTGAATTGGTGAAGAAATTTATGACCCCACTTTAAACAACTTCTGTGGCT 1701
Qy      452 AlaAlaValTTPAlaValAlaSerTyrValValAlaAspMetGluGluMetLeuPro 470
Db      1702 CAATTACGAGGACGACTGGTA---TATGAGCTTGTGAT---TCTAAATCATTTCT 1752

RESULT 9
US-10-325-430-13
; Sequence 13, Application US/10325430
; Publication No. US20030153525A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Siles-Santiago, Immaculada
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; FILE REFERENCE: MP101-294P1RNM
; CURRENT APPLICATION NUMBER: US/10/325,430
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/341,953
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 3152
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-325-430-13

Alignment Scores:
Pred. No.: 5,436-12 Length: 3152
Score: 188.00 Matches: 112
Percent Similarity: 35.24% Conservative: 85
Best Local Similarity: 20.04% Mismatches: 186
Query Match: 7.78% Indels: 176
Gaps: 27

US-09-745-763-36 (1-472) x US-10-325-430-13 (1-3152)
Qy      58 LysAlaGlnAsn-----ArgSerTyrGluArgLeuAlaLeuLeuValAsp 72
Db      211 AAAGCTGAACCATCAATCATTTCTGCTTTTACCAAGCTT----- 255
Qy      73 ThrValGlyProArgLeuSerGlySerIysAsn-----LeuGluLysAlaIleGln 89
Db      256 -----CTTCATCTGGCAGAACAGAAACAAATTTCTTGCTTCCCAAGAAATCCAA 306
Qy      90 lleMetTyrGlnAsnLeuGlnIleAspGlyLeuGluLysValIleLeuGluPro---Val 108

```

```

Db      307 ACCAGTGAAGAAATTTGACATGATTACAGCCAGAGTTGGTATATATGATGCTCTTA 366
Qy      109 ArgIleProHisTTPGluArgGlyGluGluSerAlaValMetLeuGluProArgIleHis 128
Db      367 TCTTACCCCAATGAGACAAATGCCCACTATATATGATTTGGATGAGAA-----CAT 417
Qy      129 LysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIleThrAla 148
Db      418 GAAACTGAGATTTCAAAACA---TCAATCTTAACACCACACAGATGGCTATGAGAAAT 474
Qy      149 GluValleuValIleThrSerPheAspGluLeuGlnArgAlaSerGluAlaArgGly 168
Db      475 GTTCAAAATATGTGGCCACCATATATATGCTTTCAGCCCAAGGATG---CCAGAGGA 531
Qy      169 LysIleValValTyrAsnGlnProTyrIleAsnTyrSerArgThrValGlnTyr----- 186
Db      532 GATCTTGTA-----TATGTGAATCATCTGCGACCTGAAAGACTTTTTCAAA 576
Qy      187 -----ArgThrGlnGlyValAlaGluAlaIleLysValIleGlyAla 199
Db      577 CTAGAAAGAGATGGGATCACTGATCTGGAAGATTTGATATGCAAGATATGAAAA 636
Qy      200 Leu-----AlaSerLeuIleArgSerValAlaSerPheSerIle 212
Db      637 ATCTTCAAGAGAAATAAAGTTAAATATGATGCTTACAGAGACCATAGAAATCATCTTG 696
Qy      213 TyrSer----- 214
Db      697 TACTCATCCAGCTGACTACTTGTCTGAGAGTACAGCATATCCCAAGATGAAAT 756
Qy      215 ---ProHisThrGlyIleGln----- 220
Db      757 CTCTCTGGAAGTCAAGCCCAAGAGAAATGTGTTAAATTTGAATGCTGCTGTGACCCA 816
Qy      221 -----GluTyrGlnAsp 224
Db      817 CTCACCTCAGGCTATCCAGAAAAGAAATACCTTTCAGACTTGATGTTGAAGAAGAGATG 876
Qy      225 GlyValProLysIleProThrAlaCysIleThrValGluAspAlaGluMetMetSerArg 244
Db      877 GGAATCCCCCGAAATACCTGTACATCCCATGATATATATGATGAGAAATATTTTACGC 936
Qy      245 MetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAlaLysThr----- 262
Db      937 TACTTGGGA---GGAATTCCTCCACCAATAGAGTTGGAAGGAGCCCTTAATGTAGT 993
Qy      263 -----TyrProAspThrAspSerPhe----- 269
Db      994 TATAGTATCGGACCTGGCTTTACAGGAGATGATTTCTTTCAGAGATGATGATCATGTT 1053
Qy      270 -----AsnThrValAlaGluIleThrGlySerLys 279
Db      1054 TATAACATCATATAAATTACAGAGATTTACAAATGTAGTTGAACTATCAGAGATCTGTG 1113
Qy      280 TyrProGluGlnValleuValSerGlyHisLeuAspSerTTPAspValGlyGlnGly 299
Db      1114 GAACCTGACAGATATGTATTTCTGGAGGTACCCGGAATCTCTGGATTT-----GGA 1167
Qy      300 AlaMetAspAspGlyGlyAlaPheIleSerTrpLys-----AlaLeuSerLeu 316
Db      1168 GCTATTTGACCCCAACACAGTGGGTTGCTGTTTTCAGAAATAATGCCGAGATTTTGAAAA 1227
Qy      317 IleLysAspLeuGlyLeuArgProLysArgThrIleuArgLeuValleuTrpThrAlaGlu 336
Db      1228 CTGATGAGTAAAGCTGAGACCTTGAAGAACTATATTTTTCAGACTGGATGACGAA 1287
Qy      337 GluGlnGlyGlyValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsn 356
Db      1288 GAATTTGGACTTCTGGGTTCACAGAAATGGCTGAG-----GAGAAAT 1329
Qy      357 TyrSerLeuValMetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThr 376

```

```

Db 1330 GTCAAAATCTCCAGAGAGACATT---GCTTATATCACTCGATTCTATAGAA 1386
Qy 377 GlySerGluValAlaArgAla-----11MetGluGluValMetSerLeu 391
Db 1387 GGCAATTATATCTTCAGAGTGTACTGATCCCTCTTTACCAATTAGTGTATTAAGCTG 1446
Qy 392 LeuGlnProLeu-----395
Db 1447 ACAAAAGAGATCCCGAGCCCTGATGAGTGTGAGTAATCATCTGTATGAAGCTGG 1506
Qy 396 -----Asn1LeuGlnValLeuSerHisGlyGlu 405
Db 1507 TTGGAAAAGACCCCTTCACTGAAATAATTAATTTGCTTACATCAATTAAGCTGGATCT 1566
Qy 406 GlyThrAsp1LeuAsnProTyr1LeuGln-----AlaGlyValProGluVala----- 420
Db 1567 GGAAGTGACTTTGAAGCTTATTTTCAGAGACTTGGAAATGCTTACAGCAGAGCCCGTTAC 1626
Qy 421 -----SerLeuLeuAspSerLeuTyr1LeuTyrPhePhePheHisHisSerHisGlyAsp 438
Db 1627 ACTAAGATTAAGAAACAGATTAAGTACGACAGTACCAGTACACACAAATTTATGAG 1686
Qy 439 ThrMetThrValMet-----AspPro-----LysGlnMetAsnValAla 451
Db 1687 ACATTTGAATTGGTAGAGAAATTTTATGACCCCACTTAATAAAACAATTTCTGTGCT 1746
Qy 452 AlaAlaVal1TrrAlaVal1SerTyrValValAlaAspMetGluGluMetLeuPro 470
Db 1747 CAATTACGAGAGACAGCTGTA---TATGAGCTTGTGAT---TCTAATAATCTTCT 1797

RESULT 10
US-10-252-157-9
; Sequence 9, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Paris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 415650.5
US-10-252-157-9

Alignment Scores:
Pred. No.: 8,94e-10 Length: 2655
Score: 169.00 Matches: 105
Percent Similarity: 32.65% Conservative: 71
Best Local Similarity: 19.48% Mismatches: 174
Query Match: 6.99% Indels: 189
DB: 13 Gaps: 22

US-09-745-763-36 (1-472) x US-10-252-157-9 (1-2655)
Qy 3 PheLeu1IlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCyseSerGlyValAla 22
Db 356 TTCTTTCTCCTCGGCTTCCTCTTCGGGTGTTTAAATCTCCATGAACTACTAAC 415
Qy 23 IleCysLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIleAlaSer 42
Db 416 ATTACTCCAAAGCATATATATGAAACATTTTGGATGTAATGAAAGCTGAG----- 466
Qy 43 CysGlyAspValAlaIleLysLeuLeuValTyrGlyLysAlaGlnAsnArg 62

```

```

Db 467 -----AACATCAAGAGTCTTATATTAATTT----- 493
Qy 63 SerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGlySerLys 82
Db 494 -----ACACAGATACACACTTTTACGAGAACAGAA 523
Qy 83 -----AsnLeuGluValAlaIleGln11MetTyrGlnAsnLeuGlnGlnAspGly 99
Db 524 CAAACTTCAGCTTGCAAGCAATTTCAATCCAGTGAAGAAATTTGGCCGTGATCT 583
Qy 100 LeuGluLysValHisLeuGluProVal-----Arg1LeuProHisTyr 113
Db 584 GTTGAGCTACACATTAATGATGATGCTGCTGCTTCCACCAATTAAGATCACTCCAACTAC 643
Qy 114 -----GluArgGlyGluGlu-----SerAlaValMetLeuGluProArg 126
Db 644 ATCTCAATTAATTAATGATGATGATGATGATTTTCAACATCATTAATTTGAACCA--- 700
Qy 127 IleHisLysIleAlaIleLeuGlyLeuGlySerSer1LeuGlyThrProProGluGlyIle 146
Db 701 -----CCTCCTCCAGATAT 715
Qy 147 ThrAlaGluValLeuValAlaThrSerPheAspGluLeuGlnArgArgAlaSerGluAla 166
Db 716 GAAATGTTTCGATATATGATACACACCTTCAGTCTTCTCTCCTCAAGAAATG---CCA 772
Qy 167 ArgGlyLysIleValValTyrAsnGlnProTyr11LeuAsnTyrSerArgThr----- 183
Db 773 GAGGCGATCTAGTG-----TATGTTAATCTATACGACGACACTGCAAGACTTC 817
Qy 184 -----ValGlnTyrArgThrGlnGlyValAlaGluAlaIleVal 197
Db 818 TTTAATATGGAACGGACATGAAATTAATGCTCTCGGAAATTTGATATGCCAGATAT 877
Qy 198 Gly-----AlaLeuAlaSerLeu1LeuArgSerVal 207
Db 878 GGGAAAGTTTTCAGAGGAATTAAGTTTAAATGACCCAGCTGCGAGGAGCAAGAGAGTC 937
Qy 208 AlaSerPheSer-----11LeuTyrSerProHisThrGlyIleGlnGluTyrGln 223
Db 938 ATTCTTACTCCAGACCTGCTGACTACTTGTGCTCT---GGGAGAAAGTCTTATCCA 991
Qy 224 Asp----- 224
Db 992 GATGTTGAATCTTCTCTGAGAGTGTGTCACAGCTGGAATATCTTAATCTGAATGCT 1051
Qy 224 ----- 224
Db 1052 GCAAGAGACCTCTCACACCAAGTTTACCAGCAATGATATGCTTATAGCGTGAATTT 1111
Qy 225 -----GlyValProLys11LeuProThrAlaCys11LeuThrValGluAspAlaGlu 240
Db 1112 GCAGAGCTGTGTGCTTCCAGATATTCCTGTTCAATCAATGATGATGATGATGACAG 1171
Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleVal11LeuGlnLeuLysMetGlyAla 260
Db 1172 AAGTCTCTTAAG-----AAAATGGAGTGGC 1195
Qy 261 LysThrTyrProAsp----- 265
Db 1196 TCAGACACACAGATTAACAGCTGAGAGAGAAAGTCTCAAGTGCCTCAATGTTGACCT 1255
Qy 265 ----- 265
Db 1256 GCGTTTACTGGAACCTTTTCTACACAAAAGTCAAGATGCACATCTCAATCAATGAA 1315
Qy 266 -----ThrAspSerPheAsnThrValAlaGlu11LeuThrGlySerLysTyrProGluGlnVal 284
Db 1316 GTGACAGAAATTTTACATATGATATGATGATCTCTCAGAGAGACAGTGAACACAGACATAT 1375
Qy 285 ValLeuValSerGlyHisLeuAspSerTrrPaspValGlyGlnGlyAlaMetAspAspGly 304

```

Db 1376 GTCAATTCGGAGTCACCGGACATCATGAGTGT---GTTGTAATGACCCCTCAG 1429  
 Qy 305 Gtlyglalapheliesertrpglu-----Alaleuserleuilelysaspheugly 321  
 Db 1430 AGTGAAGACCTGTTGTTTCATATAATGTGAGACCTTTGAACTGAAAGAAAGAGAG 1489  
 Qy 322 Leuagtrpolsartrleuargleuvalleuiprthralagluugluuglylval 341  
 Db 1490 TGGACACCTGAGAAACATTTGTTTGCAGCTGGATGCAAGAAATTTGCTTCTT 1549  
 Qy 342 Glialapheliegltrtyrleuuhileysvalanilleasertrserleuvalmet 361  
 Db 1550 GGTTCCTACTGAGTGGACAGTTCCTTTATATATAGAAATCAAGACTCCTTCAA 1609  
 Qy 362 Gluserapalaglythrphleuprorthrglyleuuglnphethrglyserciulylsala 381  
 Db 1610 GAGCCTGGCGTG--GCTTATATATATGCTGACTCATCTATAGAAAGAAACCTCTG 1666  
 Qy 382 Argalalilemergluvalmet---Serleuleuglnproleuasnillethrglnval 400  
 Db 1667 AGAGTTGATGTACACCCGCTGATGTACAGCTTGTACAC-----AACCTAACAAAGAG 1720  
 Qy 401 Leu---Serhisglygluglythrasp-ileasnphetrpilleglnalaglyval 417  
 Db 1721 CTGAAAGCCCTGATGAAGGCTTGAAGCAATCTCTTATGAAAGTTGACTA 1775

# RESULT 11

US-09-969-708-536  
 ; Sequence 536, Application US/09969708  
 ; Patent No. US20020102532A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Augustus, Meena  
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu  
 ; FILE REFERENCE: 689290-70  
 ; CURRENT APPLICATION NUMBER: US/09/969,708  
 ; PRIOR FILING DATE: 2001-10-03  
 ; PRIOR APPLICATION NUMBER: US/60/237,606  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: US/60/237,608  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: US/60/237,425  
 ; PRIOR FILING DATE: 2000-10-03  
 ; NUMBER OF SEQ ID NOS: 658  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 536  
 ; LENGTH: 2653  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-969-708-536

Alignment Scores:  
 Pred. No.: 2,096-09 Length: 2653  
 Score: 166.00 Matches: 105  
 Percent Similarity: 32.47% Conservative: 70  
 Best Local Similarity: 19.48% Mismatches: 169  
 Query Match: 6.87% Indels: 195  
 DB: 10 Gaps: 23

US-09-745-763-36 (1-472) x US-09-969-708-536 (1-2653)

Qy 3 Pheleuilephealaphhegllylvalhisleuuserleucyserserilylsala 22  
 Db 355 TTCTTCTCTCGGCTTCTCTCTGGGAGTGTATTAATCTCCATGAAGACTACTAC 414  
 Qy 23 ilecylsyaanglylleserlysartrthrpheglugluiletylglugluilelaser 42  
 Db 415 ATTACTCCAAAGCAATATATGAAACATTTTGTGATGAATGAAAGCTGAG----- 465  
 Qy 43 Cysglasrpvalalalyalaleileasnleuualvaltyrglylylsalaglnasnarg 62  
 Db 466 -----AACATCAAGAGTTCCTTATATATATTT----- 492

Qy 63 sertryglurgleualaleuvalaspthrvalglyproargleuserglyserlys 82  
 Db 493 -----ACAGAGTACCAKCTTTAGCAGAAACGAA 522  
 Qy 83 -----Asnleuglylsalaleglnillewertyrleuasnleuglnaspgly 99  
 Db 523 CAAACTTTCAGCTTGGAAAGCAAAATTCATCCAGTGAAGAAATTTGGCTGGAATCT 582  
 Qy 100 Leuuglylsvalhisleuglnproval-----Argilleprohiatrp 113  
 Db 583 GTTAGCTAGCACATTAATGATGTCCTGTCTCTCAACCAATTAAGCTCATCCAACTAC 642  
 Qy 114 -----Gluatrglygluglu-----seralavaleuileuglnproarg 126  
 Db 643 ATCTCAATTAATTAATGAGATGAAATGAAATTTTCAACATCATTAATTTGAACCA--- 699  
 Qy 127 ilehisyllealaleuuglyleuglyserierleeglythrproprogluglyle 146  
 Db 700 -----CCTCTCCAGATAT 714  
 Qy 147 Thrvalagluvalleuvalvalthrsrphaspoluleuglnargalalasergluala 166  
 Db 715 GAAATGTTTGGATATATGACACCTTCAAGCTTCTCTCAAGAAATG--CCA 771  
 Qy 167 Argglylysilevallyrasnglnprotyrilleasntryserargthr----- 183  
 Db 772 GAGGCCATCTAGT-----TATCTTAATCTATGACCAACCTGAAGACTTC 816  
 Qy 184 -----valglntryargthrnglylvalavallglualalalyval 197  
 Db 817 TTTAAATTTGAACCGGACATGAATAATCAATGCTCTGGAAATTTGAATTTGCCAGATAT 876  
 Qy 198 Gly-----Alaleuasnleuileargserval 207  
 Db 877 GGGAAAGTTTCAGAGAAATTAAGCTTAAATGCCAGCTGCGAGGCGCAAGAGT 936  
 Qy 208 Alaserpheser-----iletyrserprohiethrglylleglnluityrgln 223  
 Db 937 ATTCTACTCCGACCTGCTGACTACTTGTCTCT-----GGGGAAGTCTTATCCA 990  
 Qy 224 Asp----- 224  
 Db 991 GATGTTGAATCTTCTGAGAGTGTGTCCAGCGTGAATAATCTTAATCTGAATGCT 1050  
 Qy 224 ----- 224  
 Db 1051 GCAGAGACCTCTCACACAGGTTACCCAGCAATGAATATGCTTATAGCGGTGAATT 1110  
 Qy 225 -----glyvalprolysileprothrleacyllethrvalgluasplaglu 240  
 Db 1111 GCAGAGCTGTGTCTTCCAAAGTATCTCTGTCATCCATTTGATATGATGACAG 1170  
 Qy 241 Metwetserymetalaserhisglyllylsylevalilleglnleuulysmetglyala 260  
 Db 1171 AAGCTCTAGAA-----AAAATGGGTGGC 1194  
 Qy 261 lyethrlyrproasp----- 265  
 Db 1195 TCAGACACCACCATGACAGCTGAGAGAGAGTCTCAAGTCCCTTCAATGTTGACCT 1254  
 Qy 265 ----- 265  
 Db 1255 GGCCTTACTGAAACTTTTCTACACAAAAGTCAAGATGACATCCACTCTACCAATGA 1314  
 Qy 266 ---Thraserpheanthrvalalagluillethrglyserlystyrrprogluglnval 284  
 Db 1315 GTGACAGAAATTTACATGATGATGATGATCTCTGAGAGACAGAGCAACAGATAT 1374  
 Qy 285 valleuvalserglyhileuasnsertrpaspvalglynglylvalametaspsgly 304  
 Db 1375 GTCAATTTGAGAGTTCACCGGACTCATGGGTGTT-----GGTGTATTTGACCTTCAG 1428  
 Qy 305 Gtlyglialapheliesertrpglu-----Alaleuserleuilelysaspheugly 321

```

Db 1429 AGTGAAGAGCTGTTGTCATGAAATGTCGAGAGCTTTGGAACACTGAAAAGGAGG 1488
Qy 322 LeuArgProLysArgThrLeuArgLeuValLeuThrPheArgLeuGluGlyVal 341
Db 1489 TGGAGACCTTAAGAAACATTTGTTGCAAGCTGGAGTGAAGAAATTTGGCTCTCTT 1548
Qy 342 G1yAlaPheGlnTyrTyrGlnLeuHisLeuValLeuSerAsnTyrSerLeuValMet 361
Db 1549 GGTTCCTACTAGCTGGCGACAG-----GAGAAATTCAGAACTCTCTTCAA 1590
Qy 362 GluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluValAla 381
Db 1591 GAGCGTGGCGT---GCTTATATTATGCTACTCATCTATAGAAAGAACTACACTCTG 1647
Qy 382 ArgAlaIleMetGluGluValMet---SerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db 1648 AGAGTTGATTGTACACCGCTATGTACAGCTTGACAC-----ACCTTACAAAGAG 1701
Qy 401 Leu---SerHisGlyGluGlyThrAsp-IleAsnPheTrpIleGlnAlaGlyVal 417
Db 1702 CTGAAAGCCCTGATGAAAGCTTTGAAAGCAATCTCTTATGAAAGTTGAGACTA 1756

RESULT 12
US-09-873-319-451
; Sequence 451, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; APPLICANT: Kulkarni, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Waga, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OR INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873,319A
; EARLIER FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 451
; LENGTH: 2653
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 M99487
US-09-873-319-451

Alignment Scores:
Pred. No.: 2,09e-09 Length: 2653
Score: 166.00 Matches: 105
Percent Similarity: 32.47% Conservative: 70
Best Local Similarity: 19.48% Mismatches: 169
Query Match: 6.87% Indels: 195
DB: 13 Gaps: 23

US-09-745-763-36 (1-472) x US-09-873-319-451 (1-2653)
Qy 3 PheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCysSerGlyValAla 22
Db 355 TTCTTTCTCCCTCGGCTTCTCTCGGCTGTTTAAATAATCTCCAAAGAACTACTAC 414
Qy 23 IleCysLysAsnGlyIleSerLysArgThrPheGluGlnIleLysGluGlnIleAlaSer 42
Db 415 ATTACTCCAAAGCAATATATGAAAGCATTTTGTGATGATGAAAGCTGAG----- 465
Qy 43 CysGlyAspValAlaLysAlaIleIleLeuLeuAlaValIleGlyLysAlaGlnAsnArg 62
Db 466 -----AACATCAAGAGATCTTATATATTT----- 492
Qy 63 SerTyrGluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGlySerLys 82

```

```

Db 493 -----ACACAGATACCACTTATAGACGAAACGAA 522
Qy 83 -----AsnLeuGlyValAlaIleGlnIleMetTyrGlnLeuLeuGlnGlnAspGly 99
Db 523 CAAAACTTCAGCTTGAAAGCAAAATTCATCCACGGAAGAAATTTGGCTGGAATTC 582
Qy 100 LeuGluLysValHisLeuGluProVal-----ArgIleProIleTrp 113
Db 583 GTTAGCTAGACACTTATGATGTCCTTGTGCTTACCCAAATTAAGATCATCCCACTAC 642
Qy 114 -----GluArgGlyGluGlu-----SerAlaValMetLeuGluProArg 126
Db 643 ATTCATTAATTAATTAAGATGAAATGAAATGATTTTCAACATCATTAATTGAACCA--- 699
Qy 127 IleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIle 146
Db 700 -----CCTCTCCAGGATAT 714
Qy 147 ThrAlaGluValLeuValValThrSerPheAspGluLeuGlnArgArgAlaSerGluAla 166
Db 715 GAAATGTTCCGATATGTGACACCTTCAGTCTTCTCTCTCAAGAAATG---CCA 771
Qy 167 ArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyrSerArgThr----- 183
Db 772 GAGGCGCATCTAGT-----TATGTTAACTATGACCACTGAAGACTTC 816
Qy 184 -----ValGlnTyrArgThrGlnGlyValAlaGluAlaAlaLysVal 197
Db 817 TTTAAATGGAACGGACATGAATCAATGCTCTCGGAAATGTAATGTAATGCCAGATAT 876
Qy 198 Gly-----AlaLeuLaserLeuIleArgSerVal 207
Db 877 GGGAAAGTTTTCACAGAAATTAAGCTTAAATGCCCCAGCTGGAGGGCCAAAGAGTTC 936
Qy 208 AlaSerPheSer-----IleTyrSerProHisThrGlyIleGlnGluTyrGln 223
Db 937 ATTCTACTCCACCGCTGCTGACTACTTGTCTCT-----GGGGAAGTCTTATCCA 990
Qy 224 Asp----- 224
Db 991 GATGTTGGAATCTTCTCGAGAGTGTCGTCACAGGTGAAATATCTTAATCTGAATGCT 1050
Qy 224 ----- 224
Db 1051 GCAGAGACCTCTCACACCAAGTTACCCAGCAATATGCTTATAGCGGTGAATT 1110
Qy 225 -----GlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db 1111 GCAGAGCTGTGCTCTTCCAGATATCTGTTCAATCATTCATGATATGATGACACAG 1170
Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260
Db 1171 AAGCTCTTAGA-----AAAATGGGTGGC 1194
Qy 261 LysThrTyrProAsp----- 265
Db 1195 TCAGCACCAACCATGATGACAGCTGAGAGAACTCAAAAGTCCCTCAATGTTGACCT 1254
Qy 265 ----- 265
Db 1255 GCGTTTACTGGAACCTTTCTACACAAAAAGTCAAGATGACATCCACTACCAATGAA 1314
Qy 266 ---ThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyrProGluGlnVal 284
Db 1315 GTGACAAAGAAATTTACATATGATGATGATCTCGACAGACAGTGAACAGACAGATAT 1374
Qy 285 ValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAlaMetAspAspGly 304
Db 1375 GTCAATTCGGAGGTCAACCGGACTCAAGGATGTT-----GGTGGTATTTGACCTTCAG 1428
Qy 305 G1yGlyAlaPheIleSerTrpGlu-----AlaLeuSerLeuIleLysAspLeuGly 321

```

Qy	83	-----	AsnLeuGluValAlaIleGlnIleMet	ArgIleProHis	199	
Db	523	CAAACTTCACCTTGCAAGCAAAATTAATCCAGTGGAAAGAAATTTGGCCGTGGAAATTC	582	-----		
Qy	100	LeuGluValAlaIleLeuGluProVal	-----	ArgIleProHis	113	
Db	583	GTTGACCTAGCACTTATGATGTCCTGTCGTCCTACCAAAATGAAGTACTCCCACTAC	642	-----		
Qy	114	-----	GluArgGlyGluGlu	-----	SerAlaValMetLeuGluProArg	126
Db	643	ATCTCAATTAATTAAGATGGAATGGAATGATTTTCAACACATCATTAATTTGAACCA	699	-----		
Qy	127	IleHisLeuAlaIleLeuGlyLeuGlySerIleGlyThr	ProProGluGlyIle	146		
Db	700	-----	-----	CCTCTCCAGAAAT	714	
Qy	147	ThrAlaGluValLeuValValThrSerPheAspGluLeuGlnArgAlaSerGluAla	166	-----		
Db	715	GAAATAGTTTGGATATTTGTAACACCTTCAGTGCCTTCTCTCCTCAAGAAAG	771	-----	CCA	
Qy	167	ArgGlyLeuAlaValValThrAsnGlnProTyrlleAsnTySerArgThr	-----	183		
Db	772	GAGGGGAGATCTAGTG	-----	TATGTTAACTATGACACAGACATGAAAGACTTC	816	
Qy	184	-----	ValGlnTyArgThrGlnGlyAlaValGluAlaAlaValVal	197		
Db	817	TTTAAATTTGAAACGGACACATGAATAATCAATTCCTGGAAATTTGTAATTCGCAATAT	876	-----		
Qy	198	Gly	-----	AlaLeuAlaSerLeuIleArgSerVal	207	
Db	877	GGGAAAGTTTTCAGAGCAATTAAGTTTAAATATCCACGCTGGACGGGCGCAAGAGATC	936	-----		
Qy	208	AlaSerPheSer	-----	IleTySerProHisThrGlyIleGlnIleTyGln	223	
Db	937	ATCTCTACTCCGACCTCGTGACTACTACTTCTCTCT	-----	GGGCGGAAGTCTTATCCA	990	
Qy	224	Asp	-----	224		
Db	991	GATGTTTGAATCTTCTCGAGGTGGTGTCCAGCGTGGAAATATCTTAAATCTGAATGCT	105	-----		
Qy	224	-----	224	-----		
Db	1051	GCAGAGACCTCTCACACACAGGTTACCCACACAAATGAATATGCTTATAGCGTGAAT	111	-----		
Qy	225	-----	GlyValProIleIleProThrAlaCysIleThrValGluAspAlaGlu	240		
Db	1111	GCAGAGCTGTGGTCTTCCAAAGTATCTCTGTCATCCAAATGGATCTATGATGACAG	117	-----		
Qy	241	MetMetSerArgMetAlaSerHisbelyIleValIleGlnLeuIleMetGlyAla	260	-----		
Db	1171	AAGCTCTTAGAA	-----	AAAAATGGGTGC	119	
Qy	261	LysThrTyProAsp	-----	265		
Db	1195	TCAGCACCAACAGATAGACGCTGAGAGGAAGTCTCAAAGTCCCTTCAATGTTGACCT	125	-----		
Qy	265	-----	265	-----		
Db	1255	GGCTTACTGAAACCTTTCTACACAAATAAGTCAAGATGACATCCACTTACCAATGAA	131	-----		
Qy	266	-----	ThrAspSerPheAsnThrValAlaGluIleThrIleTySerTyProGluGlnVal	284		
Db	1315	GTCACAAAGATTTTCAATATGATAGTACTCTTCAGAGACACAGATGAAACAGACAGATAT	137	-----		
Qy	285	ValLeuValSerGlyHisLeuAspSerThrAspArgAlaGlyGlnGlyAlaMetAspArgGly	304	-----		
Db	1375	GTCATCTCGGAGGTACCCGGGACTCATGGGTGTTT	-----	GGGTGATATGACCTCAG	142	
Qy	305	GlyGlyAlaPheIleSerTrpGlu	-----	AlaLeuSerLeuIleTyAspLeuGly	321	
Db	1429	AGTGGAGACAGCTGTTGTCATGAATATGTGACAGAGCTTTGGAAACATCGAAAAAGAAAGG	148	-----		
Qy	322	LeuArgProLysArgThrLeuArgLeuValLeuThrThrAlaGluGluGlnGlyVal	341	-----		



Db 1489 TGGAGACCTGACAAACAAATTTTGTTCAGAGCTGGAGATGCAGAAAGATTTGGCTTCTT 1548  
Qy 342 GYAlaIlePheGlnIYrTYrGlnLeuHieIyValAsnIleSerAsnTYrSerLeuValMet 361  
Db 1549 GGTTCCTACTGATGGGACAGAG-----GAGAAATTCAGAGCTCTCCAA 1590  
Qy 362 GluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyValAla 381  
Db 1591 GAGGCTGGCCGCG---GCTATATTAAATGCTGATCATCATATAGAAAGAACTACACTCTG 1647  
Qy 382 ArgAlaIleMetGluGluValMet---SerLeuGlnProLeuAsnIleThrGlnVal 400  
Db 1648 AGAGTTGATTGTACACCGCTGATGTACAGCTTGATCAC-----AACCTAACAAGAG 1701  
Qy 401 Leu---SerHisGlyGluGlyThrAsp-IleAsnPheThrIleGlnAlaGlyVal 417  
Db 1702 CTGAAAGCCTGATGAAAGGCTTGAAAGCAATCTTTATGAAAGTTGAGACTA 1756

## RESULT 14

US-10-210-120-5  
Sequence 5, Application US/10210120  
Publication No. US20030175736A1  
GENERAL INFORMATION:  
APPLICANT: Chinnaiyan, Arul M.  
APPLICANT: Rubin, Mark A.  
APPLICANT: Sreekumar, Arun  
TITLE OF INVENTION: Expression Profile of Prostate Cancer  
FILE REFERENCE: UM-07221  
CURRENT APPLICATION NUMBER: US/10/210,120  
CURRENT FILING DATE: 2002-08-01  
PRIOR APPLICATION NUMBER: US 60/309,581  
PRIOR FILING DATE: 2001-08-02  
PRIOR APPLICATION NUMBER: US 60/334,468  
PRIOR FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 123  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 5  
LENGTH: 2653  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-210-120-5

## Alignment Scores:

Pred. No.: 2,096-09 Length: 2653  
Score: 166.00 Matches: 105  
Percent Similarity: 32.47% Conservatve: 70  
Best Local Similarity: 19.48% Mismatches: 169  
Query Match: 6.87% Indels: 195  
Gaps: 23

US-09-745-763-36 (1-472) x US-10-210-120-5 (1-2653)

Qy 3 PheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGlyValAsnAla 22  
Db 355 TTCTTTCTCCTCGGCTTCTCTGCGGTGTTTAAATCTCCAAATGAGACTACTAC 414  
Qy 23 IleCySerAsnGlyIleSerIleYrGlyThrPheGluGluIleYrGluGluIleAlaSer 42  
Db 415 ATTACTCCAAAGCAATATATGAAAGCATTTTGGATGATGAAAGCTGAG----- 465  
Qy 43 CysGlyAspValAlaIleValAlaIleLeuLeuAlaValTYrGlyValAlaGlnAsnArg 62  
Db 466 -----AACATCAAGAACTTCTATATATATTT----- 492  
Qy 63 SerTYrGluArgLeuAlaLeuValAspThrValGlyProArgLeuSerGlySerIys 82  
Db 493 -----ACACGATACCACTTTCAGCAAGAAAGAA 522  
Qy 83 -----AsnLeuGluValAlaIleGlnIleMetTYrGlnAsnLeuGlnGlnAspGly 99  
Db 523 CAAAACTTTCAGCTTGCAAGCAAAATTCATCCCGAGGAAAGAAATTTGGCTGATTTCT 582

Qy 100 LeuGlyValValHisLeuGluProVal-----ArgIleProHisTrp 113  
Db 583 GTTAGCTACACATTAATGATGTCCTGTCTCCACCAATAAGACTATCCCACTAC 642  
Qy 114 -----GluArgGlyGluGlu-----SerAlaValMetLeuGluProArg 126  
Db 643 ATCTCAATTAATTAATGAAAGATGAAATGAGATTTTCACACATCATTAATTTGAACCA--- 699  
Qy 127 IleHisIysIleAlaIleLeuGlyLeuGlySerSerIleGlyThrProProGluGlyIle 146  
Db 700 -----CCTCCTCCAGGATAT 714  
Qy 147 ThrAlaGluValLeuValValThrSerPheAspGluLeuGlnArgArgAlaSerGluAla 166  
Db 715 GAAATGTTCCGATATGTATGACCACTTACAGGCTTCTCTCTCCACAGGAATG---CCA 771  
Qy 167 ArgGlyLysIleValValTYrAsnGlnProTYrIleAsnTYrSerArgThr----- 183  
Db 772 GAGGCGCATCTAGT-----TATGTTAACTATGACGACGACGAAAGACTTC 816  
Qy 184 -----ValGlnTYrArgThrGlnGlyAlaValAlaGluAlaIleVal 197  
Db 817 TTTAATGGAAGGACATGAAATCAATTCCTCTGGAAATGTAAATGCGAGATAT 876  
Qy 198 Gly-----AlaLeuAlaSerLeuIleArgSerVal 207  
Db 877 GGGAAATTTTCAGAGAAATTAAGTTAAATATGCCAGCTGCAGGGCCAAAGAGTGC 936  
Qy 208 AlasPheSer-----IleTYrSerProHisThrGlyIleGlnIleTYrGln 223  
Db 937 ATTCTACTCCAGCCCTGCTGACTACTTGTCTCT-----GGGAGAGTCTATGCCA 990  
Qy 224 Asp----- 224  
Db 991 GATGTTGGAATCTTCTGAGAGTGCTGCCAGCGTGAATATCTAAATTCGAATGT 1050  
Qy 224 ----- 224  
Db 1051 GCAGAGACCCCTCTACACACAGGTTACCCAGCAATGATATGTTATAGCGTGAAT 1110  
Qy 225 -----GlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
Db 1111 GCAGAGCTGTGGTCTTCCAGATATCTGTTCATCATCAATGGATACTATGATGCACAG 1170  
Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
Db 1171 AAGCTCTAGAA-----AAAATGGGTGGC 1194  
Qy 261 LysThrTYrProAsp----- 265  
Db 1195 TCAGCACCAACGATAGACAGTGAAGAGAAAGTCTCAAAGTCCCTCAATGTGGACCT 1254  
Qy 265 ----- 265  
Db 1255 GGCTTACTGAAAATTTTCTACACAAAAGTCAAGATGCATCTACCAATGAA 1314  
Qy 266 -----ThrAspSerPheAlaThrValAlaGluIleThrGlySerIYrTYrProGluGlnVal 284  
Db 1315 GTGACAAAGAAATTAATGATGATAGTACTCAGAGAGACAGTGAACAGACAGATAT 1374  
Qy 285 ValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAlaMetAspAspGly 304  
Db 1375 GTCAATTTGGGAGGTCCCGGACTCATGGGTGTT-----GGTGGTATTTAGCCCTCAG 1428  
Qy 305 GlyGlyAlaPheIleSerTrpGlu-----AlaLeuSerLeuIleLysAspLeuGly 321  
Db 1429 AGTGAGCAGCTGTGTTCATGAATGTGAGAGGCTTGGAACACTGAAAGAAAGAGGG 1488  
Qy 322 LeuArgProLysArgThrIleuValArgLeuValLeuTrpThrAlaGluGluGlnGlyVal 341  
Db 1489 TGGAGACCTTAGAGAAATTTTGTGCAAGCTGGGATCCAGAGAAATTTGGCTTCTT 1548  
Qy 342 GYAlaIlePheGlnIYrTYrGlnLeuHieIyValAsnIleSerAsnTYrSerLeuValMet 361

```

Db      1549 GGTCTACTAGTGGGCAAG-----GAGAAATTCAGACTCTTCA 1590
Qy      362 GUSERAEPALAGLThrheleuprothrglyleuglnphethrglysergluysala 381
Db      1591 GAGCGTGCGGNT---GCTATATATATAGCTGACTATCTTAGAAGAACTACACTCTG 1647
Qy      382 ARGALAIleMetgluValMet---SerleuLeuGlnProleuAsnIleThrglnVal 400
Db      1648 AGAGTTGATGTACACCGCTGATGACCTTGTRACAC-----AACCTAACAAAGAG 1701
Qy      401 Ieu---SerHieglYgluGlyThrasp-IleasnPhetPrlleglnalaglyVal 417
Db      1702 CTGAAGAGCCCTGATGAGGCTTTGAAAGCAATCTTTATGAAAGTTGACTA 1756

RESULT 15
US-10-341-434-70
/ Sequence 70, Application US/10341434
/ Publication No. US20030215835A1
/ GENERAL INFORMATION:
/ APPLICANT: Origene Technologies
/ TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
/ FILE REFERENCE: 9U 204 205 R1
/ CURRENT APPLICATION NUMBER: US/10/341,434
/ PRIOR FILING DATE: 2003-07-18
/ PRIOR APPLICATION NUMBER: US 60/348,164
/ PRIOR FILING DATE: 2002-01-15
/ PRIOR APPLICATION NUMBER: US 60/348,119
/ NUMBER OF SEQ ID NOS: 238
/ SOFTWARE: Patent version 3.1
/ SEQ ID NO 70
/ LENGTH: 2653
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (262..(2511)
/ OTHER INFORMATION:
US-10-341-434-70

Alignment Scores:
Pred. No.: 2,09e-09 Length: 2653
Score: 166.00 Matches: 105
Percent Similarity: 32.47% Conservative: 70
Best Local Similarity: 19.48% Mismatches: 169
Query Match: 6.87% Indels: 195
DB: 13 Gaps: 23

US-09-745-763-36 (1-472) x US-10-341-434-70 (1-2653)
Qy      3 PheleullePheAlPhePheglYglYValHIsleuLeuSerLeuCySerGlyYsala 22
Db      355 TTCTTCTTCCTCGGCTTCTCTTCGCGGTGTTATATAATCTCCAAATGAACTACTAC 414
Qy      23 IleCylysAsnGlyYIleSerlysaRgThrPhegluGluIlelyleuglnuIleaser 42
Db      415 ATTACTCCAAAGCAATATATGAAAGCATTTTGGATGATGAAAGCTGAG----- 465
Qy      43 CySglYAspValAlIleYsAlIleIleAsnLeuAlaValYrglyYleAlaGlnAsnArg 62
Db      466 -----AACATCAAGAGTTCTTATATAATTT----- 492
Qy      63 SerYrgluYrgleuAlaIleuValAspThrValGlyProArgLeuSerGlySerlyS 82
Db      493 -----ACACAGATACCACTTTCAGAGAACAA 522
Qy      83 -----AsnleuGluYsAlIleGlnIleMetYrglnAsnleuGlnGlnAspGly 99
Db      523 CAAACTTTCAGCTTCGAAAGCAATTCATATCCAGTGAAGAAATTTGGCTGATTCCT 582
Qy      100 leuGluYsValHIsleuGlnProVal-----ArgIleProHietrP 113

```

```

Db      583 GTTGAGCTAGACATTAATGATGCTGTGCTTACCCAAATTAAGTATCCCACTAC 642
Qy      114 -----GluArgYgluGlu-----SerAlaValMetLeuGluProArg 126
Db      643 ATTCATATATTAATGAAGATGAAATGATTTTCAACATCATATTAATTTGAACCA-- 699
Qy      127 IleHIsyAlleAlaIleleuGlySerSerIleGlyThrProProGluGlyIle 146
Db      700 -----CCTCCCTCAAGATAT 714
Qy      147 ThrAlaGluValIleuValValThrSerPheAspGluLeuGlnArgAlaSerGluAla 166
Db      715 GAAATGTTTGGATATGTATGACACCTTCAGTCTCTCTCCCAAGAAATG---CCA 771
Qy      167 ArgGlyYsIleValIleValYrAsnGlnProbyrIleAsnYrSerArgThr----- 183
Db      772 GAGGCGATCTAGTG-----TATGTAACTATGCAAGAACTGAAGACTTC 816
Qy      184 -----ValGlnYrArgThrglnGlyAlaValAlaIleVal 197
Db      817 TTATAATTTGAACGGACATGAATATCATTCCTGGAAATTTGAAATTTGCCAGATAT 876
Qy      198 Gly-----AlaIleuAlaSerLeuIleArgSerVal 207
Db      877 GCGAAAGTTTTCAGAGAAATTAAGTTAAATCCCAAGCTGCGAGGGGCCAAAGAGTC 936
Qy      208 AlaSerPheSer-----IleYrSerProHietrHrglyIleGlnGluYrgln 223
Db      937 ATTCTTACTCCGACCTGCTGACTGACTGTGCTCT-----GGGATGAAGTCTTATCCA 990
Qy      224 Asp----- 224
Db      991 GATGGTTGAATCTTCTCGAGAGTGCTGCCAGGTGGAATATCTTAATCTGAATGT 1050
Qy      224 ----- 224
Db      1051 GCAGAGACCTCTCACACAGTTACCAAGTAATGCTTATAGCGGTGAATTT 1110
Qy      225 -----GlyValProlyIleProthrlaCySleIleThrValGluAspAlaGlu 240
Db      1111 GCAGAGCTGTGGTCTTCCTCAAGATATTCCTGTTATCCAAATGATATGATATGCAAG 1170
Qy      241 MetMetSerArgMetAlaSerHieglYIleYsIleValIleGlnLeuYMetGlyAla 260
Db      1171 AAGCTCTGAA-----AAATGGGTGCC 1194
Qy      261 LysThrYrProAsp----- 265
Db      1195 TCAGCACACACAGATACAGCTGAGAGAAAGTCAAGTCCCTCAATGTTGACCT 1254
Qy      265 ----- 265
Db      1255 GCGTTTACTGAAACTTTTCTACACAAAGTCAAGATGCATCTCAATCCAAATGAA 1314
Qy      266 ---ThraSerPheAsnThrValAlaGluIleThrglySerlySyrProGluGlnVal 284
Db      1315 GTGACAGAAATTTCAATGATGATGATGATCTCAGAGAGACAGTGAACAGAGATAT 1374
Qy      285 ValIleuValSerGlyYHIsleuAspSerTrpAspValGlyGlnGlyAlaMetAspAspGly 304
Db      1375 GTCAATTCGGAGGCTCACCGGACTCATGGGTGTT-----GGTGTATTAAGCCCTCAG 1428
Qy      305 GlyYglYAlaPheIleSerTrpGlu-----AlaIleuSerLeuIleYAspLeuGly 321
Db      1429 AGTGAAGACAGTGTGTTTCATGAATTTGTAGAGAGCTTTGCAACATGAAAAGAAAGG 1488
Qy      322 LeuArgProlysaRgThrLeuArgLeuValIleuThrHraIleGlnGlnGlyVal 341
Db      1489 TGGAGACCTTAAGAAACAAATTTTGTGCAAGCTGGAGATGCAAGAAATTTGGCTTCTT 1548
Qy      342 GlyAlaPheGlnYrgYrglnIleuHIsleuValAsnIleSerAsnYrSerLeuValMet 361
Db      1549 GGTCTTACTAGTGGGCAAG-----GAGAAATTCAGACTCTTCA 1590

```

```
Qy      362 GluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLysAla 381
          |||
          ::::
Db      1591 GAGCGTGGCCG---GCTTATATTAAAGCTGACTCATCTATAGAGAACTACACTCTG 1647
          |||
Qy      382 ArgAlaIleMetGluGluValMet--SerLeuLeuGlnProLeuAsnIleThrGluVal 400
          |||
          ::::
Db      1648 AGAGTTGATTGTACACCGCTGATGTACAGCTTGGTACAC-----AACCTAACAAAGAG 1701
          |||
Qy      401 Leu--SerHisGlyGluGlyThrAsp-IleAsnPheTrpIleGlnAlaGlyVal 417
          |||
          ::::
Db      1702 CTGAAAGCCCTGATGAGAGGCTTGAAGCAAAATCTCTTATGAAAGTTGACTA 1756
          |||
```

Search completed: December 23, 2003, 01:53:39  
Job time : 435 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:53:06 ; Search time 35 Seconds  
(without alignments)  
73.729 Million cell updates/sec

Title: US-09-745-763-36\_COPY\_231\_240  
Perfect score: 10  
Sequence: 1 TACTIVEDAE 10

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_plant:\*  
10: sp\_rodent:\*  
11: sp\_virus:\*  
12: sp\_vertebrate:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	100.0	433	11	070216
2	10	100.0	470	11	09WVJ3
3	10	100.0	472	4	08NBZ1
4	10	100.0	472	4	09Y646
5	10	100.0	541	4	09Y5X6
6	7	70.0	576	11	08BGT9
7	6	60.0	157	10	09SWJ1
8	6	60.0	167	2	09L990
9	6	60.0	171	16	09KMC6
10	6	60.0	185	2	08KVK6
11	6	60.0	186	2	08KVC0
12	6	60.0	186	2	08KVG2
13	6	60.0	186	2	08KHQ8
14	6	60.0	186	13	09PSY2
15	6	60.0	187	16	09ABZ8
16	6	60.0	188	11	08K4X9

17	6	60.0	197	16	09KX80	09km80 vibrio chol
18	6	60.0	205	11	09LXU1	09ixul rattus norv
19	6	60.0	224	4	09UFP5	09ufes homo sapien
20	6	60.0	220	4	09S5S3	09S5S3 homo sapien
21	6	60.0	246	4	09S5S2	09S5S2 homo sapien
22	6	60.0	260	2	052442	052442 pseudomonas
23	6	60.0	261	2	09Z158	09Z158 pseudomonas
24	6	60.0	271	11	09RZAS	09RZAS mus musculu
25	6	60.0	281	16	08PUZ8	08PUZ8 xanthomonas
26	6	60.0	293	4	09E6J3	09E6J3 homo sapien
27	6	60.0	299	11	09CW34	09CW34 mus musculu
28	6	60.0	309	17	08PYR5	08PYR5 methanosarc
29	6	60.0	315	4	09P0X8	09P0X8 homo sapien
30	6	60.0	319	4	096PU6	096PU6 homo sapien
31	6	60.0	319	4	096PU7	096PU7 homo sapien
32	6	60.0	319	11	09QUH4	09QUH4 mus musculu
33	6	60.0	321	4	09P0Y1	09P0Y1 homo sapien
34	6	60.0	323	4	096KA3	096KA3 homo sapien
35	6	60.0	325	4	0969L9	0969L9 homo sapien
36	6	60.0	325	11	061110	061110 mus musculu
37	6	60.0	328	16	08G3H1	08G3H1 bifidobacte
38	6	60.0	330	2	09KM44	09KM44 wolbachia s
39	6	60.0	330	2	09KM37	09KM37 wolbachia s
40	6	60.0	330	11	08K4Y0	08K4Y0 mus musculu
41	6	60.0	332	2	08RPM0	08RPM0 ehrlichia c
42	6	60.0	332	13	09YH19	09YH19 gallus gall
43	6	60.0	333	16	0910X1	0910X1 pseudomonas
44	6	60.0	334	2	08GLV8	08GLV8 ticketstia
45	6	60.0	334	16	09Z1N0	09Z1N0 ticketstia

## ALIGNMENTS

RESULT 1  
ID 070216 PRELIMINARY; PRT; 433 AA.  
AC 070216;  
DT 01-AUG-1998 (TREMBLER1.07, Created)  
DT 01-AUG-1998 (TREMBLER1.07, Last sequence update)  
DT 01-MAR-2003 (TREMBLER1.23, Last annotation update)  
DE Hematopoietic lineage switch 2.  
GN PGCP OR HLS2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Williams J.H., Chan C.-Y., Klinken S.P.;  
RT "Hematopoietic lineage Switch 2 (HLS2), a novel mRNA species induced during an erythroid to myeloid lineage switch."  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF009513; AAC17945.1; -  
DR MEROPS; M28.014; -  
DR MCD; MGI:1869205; Pgcp.  
DR InterPro; IPR001464; Annexin.  
DR InterPro; IPR003137; PA.  
DR PROSITE; PS00223; ANNEXIN; 1.  
DR PROSITE; PS50840; PA; 1.  
SQ SEQUENCE 433 AA; 47858 MW; 89C9B93712004669 CRC64;  
Query Match 100.0%; Score 10; DB 11; Length 433;  
Best Local Similarity 100.0%; Pred. No. 0.0031;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TACTIVEDAE 10  
|||  
Db 229 TACTIVEDAE 238

RESULT 2  
Q9WVJ3

```

ID Q9WVJ3 PRELIMINARY; PRT; 470 AA.
AC Q9WVJ3:
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DE 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Aminoacylase (Similar to plasma glutamate carboxypeptidase).
GN PCGP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Liu C.H., Lin B.Y., Chang L.Y.;
RT "Cloning of the mouse aminopeptidase gene."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strauberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body, and Mesonephros;
RX MEDLINE=22354683; PubMed=12466851;
RA The RIKEN Genome Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AF107835; AAD43215.1; -
DR EMBL; BC037067; AAH32105.1; -
DR EMBL; AK075686; BAC35891.1; -
DR MEROPS; M28.014; -
DR MGD; MGI:1889205; PgcP.
DR InterPro; IPR001464; Annexin.
DR InterPro; IPR003137; PA.
DR PROSITE; PS00223; ANNEXIN; 1.
DR PROSITE; PS50840; PA; 1.
DR Carboxypeptidase.
SQ SEQUENCE 470 AA; 51813 MW; 0F3490681691866A CRC64;

Query Match 100.0%; Score 10; DB 11; Length 470;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACTIVEDAE 10
DB 229 TACTIVEDAE 238

RESULT 3
Q9NBZ1 PRELIMINARY; PRT; 472 AA.
ID Q9NBZ1:
AC Q9NBZ1:
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ90651.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hiro Y., Saito K.,
RA Yamamoto Y., Wakamatsu A., Nakamura Y., Kojima S., Nagahashi K.,
RA Masuko Y., Ota T., Okano K., Yoshikawa Y., Aotoku S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;

```

```

RT "NEBO human cDNA sequencing project."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK075132; BAC11423.1; -
DR InterPro; IPR003137; PA.
DR PROSITE; PS50840; PA; 1.
KW Hypothetical protein.
SQ SEQUENCE 472 AA; 51784 MW; 83F2596F5D5C2F3 CRC64;

Query Match 100.0%; Score 10; DB 4; Length 472;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACTIVEDAE 10
DB 231 TACTIVEDAE 240

RESULT 4
Q9Y646 PRELIMINARY; PRT; 472 AA.
ID Q9Y646:
AC Q9Y646:
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DE 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Aminoacylase (Plasma glutamate carboxypeptidase).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Liu C.H., Lin B.Y., Chang L.Y.;
RT "Cloning of the human aminopeptidase gene."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strauberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF107834; AAD43214.1; -
DR EMBL; BC020689; AAD20689.1; -
DR MEROPS; M28.014; -
DR InterPro; IPR003137; PA.
DR PROSITE; PS50840; PA; 1.
DR Carboxypeptidase.
SQ SEQUENCE 472 AA; 51887 MW; EB6CBD2149E042BF CRC64;

Query Match 100.0%; Score 10; DB 4; Length 472;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACTIVEDAE 10
DB 231 TACTIVEDAE 240

RESULT 5
Q9Y5X6 PRELIMINARY; PRT; 541 AA.
ID Q9Y5X6:
AC Q9Y5X6:
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DE 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Blood plasma glutamate carboxypeptidase precursor (EC 3.4.17.21).
GN PCGP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99223495; PubMed=10206990;

```

RA Gingras R., Richard C., El-Alfy M., Morales C.R., Potier M.,  
 RA Feheretsky A.V.;  
 RT "Purification, cDNA cloning, and expression of a new human blood  
 RT plasma glutamate carboxypeptidase homologous to N-acetyl-aspartyl-  
 RT alpha-glutamate carboxypeptidase/prostate-specific membrane antigen."  
 RL J. Biol. Chem. 274:11742-11750 (1999).  
 DR EMBL; AF119386; AAD31418.1; -.  
 DR MEROPS; M28.014; -.  
 DR InterPro; IPR003137; PA.  
 DR PROSITE; PS50840; PA; 1.  
 KM Carboxypeptidase; Hydrolase; Signal.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 45 541 BLOOD PLASMA GLUTAMATE CARBOXYPEPTIDASE.  
 SQ SEQUENCE 541 AA; 59931 MW; 9B138FF897081F6B CRC64;  
 QY Query Match 100.0%; Score 10; DB 4; Length 541;  
 Best Local Similarity 100.0%; Pred. No. 0.0037;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 231 TRACITVEDAE 240  
 RESULT 6  
 Q8BGT9 PRELIMINARY; PRT; 576 AA.  
 AC Q8BGT9;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical glycosyl transferase.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Cecum, and Thymus;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium.  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573 (2002).  
 DR EMBL; AK033638; BAC28401.1; -.  
 DR EMBL; AK042133; BAC31179.1; -.  
 KM Hypothetical protein.  
 SQ SEQUENCE 576 AA; 66540 MW; 706310A99FCDACD5 CRC64;  
 QY Query Match 70.0%; Score 7; DB 11; Length 576;  
 Best Local Similarity 100.0%; Pred. No. 7.1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 500 ACTIVED 506  
 RESULT 7  
 Q9SMW31 PRELIMINARY; PRT; 157 AA.  
 AC Q9SMW31;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical 18.1 kDa protein.  
 GN F13M23.90 OR ATAG24950.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC euroside II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;

RN [1]  
 RA SEQUENCE FROM N.A.  
 RP Bevan M., Wedler H., Wedler E., Wambutt R., Hohnsels J., Mewes H.W.,  
 RA Mayer K.F.X., Scheller C.;  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RA SEQUENCE FROM N.A.  
 RP EU Arabidopsis sequencing project;  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RA SEQUENCE FROM N.A.  
 RP Wedler H., Wedler E., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL035523; CAB36737.1; -.  
 DR EMBL; AL161562; CAB79404.1; -.  
 KM Hypothetical protein.  
 SQ SEQUENCE 157 AA; 18115 MW; 72AFC01DA157BC3D CRC64;  
 QY Query Match 60.0%; Score 6; DB 10; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 39 ITVEDA 44  
 RESULT 8  
 Q9L990 PRELIMINARY; PRT; 167 AA.  
 AC Q9L990;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Vco28.  
 GN VCO28.  
 OS Vibrrio cholerae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrrio.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=569b;  
 RX MEDLINE=20476433; PubMed=11021935;  
 RA Clark C.A., Purins L., Kaewrakon P., Focareta T., Manning P.A.;  
 RT "The Vibrrio cholerae O1 chromosomal integrin."  
 RL Microbiology 146:2605-2612 (2000).  
 DR EMBL; AF179596; AAF71183.1; -.  
 DR InterPro; IPR000182; GCSNacetyltransf.  
 DR Pfam; PF00583; Acetyltransf. 1.  
 SQ SEQUENCE 167 AA; 19002 MW; 559971F39B1D02F9 CRC64;  
 QY Query Match 60.0%; Score 6; DB 2; Length 167;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 5 TVEDAE 10  
 Db 7 TVEDAE 12  
 RESULT 9  
 Q9KMC6 PRELIMINARY; PRT; 171 AA.  
 AC Q9KMC6;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Acetyltransferase, putative.

```

GN VCA0436.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Teitelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Ueberlack T., Fleischmann R.D., Niernan W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL: AE004377; AAF96342.1; -.
DR TIGR: VCA0436; -.
DR InterPro: IPR000182; GCN5acetyltransf.
DR Pfam: PF00583; Acetyltransf. 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 171 AA; 19504 MW; CB18644753E0F834 CRC64;

Query Match 60.0%; Score 6; DB 16; Length 171;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ITVEDA 10
Db 11 ITVEDA 16

RESULT 10
Q8KV6 PRELIMINARY; PRT; 185 AA.
ID Q8KV6
AC Q8KV6;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE 60 kDa chaparonin (Fragment).
GN CPN60.
OS uncultured pig faeces bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=190966;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22034982; PubMed=12039767;
RA Hill J.E., Seipp R.P., Betts M., Hawkins L., Van Kessel A.G.,
RA Crosby W.L., Hemmingsen S.M.;
RT "Extensive profiling of a complex microbial community by high-
RT throughput sequencing."
RL Appl. Environ. Microbiol. 68:3055-3066(2002).
DR EMBL: AF436976; AAM49237.1; -.
DR InterPro: IPR002423; Cpn60/TCP-1.
DR Pfam: PF00118; cpn60_TCP1; 1.
KW ATP-binding; Chaparone.
FT NON TER 1
FT NON TER 185
SQ SEQUENCE 185 AA; 20485 MW; E9A9FC0397C5D322 CRC64;

Query Match 60.0%; Score 6; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
Db 85 ITVEDA 90

RESULT 11

```

```

Q8KVCO PRELIMINARY; PRT; 186 AA.
ID Q8KVCO
AC Q8KVCO;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE 60 kDa chaparonin (Fragment).
GN CPN60.
OS uncultured pig faeces bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=190966;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22034982; PubMed=12039767;
RA Hill J.E., Seipp R.P., Betts M., Hawkins L., Van Kessel A.G.,
RA Crosby W.L., Hemmingsen S.M.;
RT "Extensive profiling of a complex microbial community by high-
RT throughput sequencing."
RL Appl. Environ. Microbiol. 68:3055-3066(2002).
DR EMBL: AF437097; AAM49358.1; -.
DR InterPro: IPR002423; Cpn60/TCP-1.
DR Pfam: PF00118; cpn60_TCP1; 1.
KW ATP-binding; Chaparone.
FT NON TER 1
FT NON TER 186
SQ SEQUENCE 186 AA; 20188 MW; D798608CB07C9E7C CRC64;

Query Match 60.0%; Score 6; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
Db 85 ITVEDA 90

RESULT 12
Q8KV2 PRELIMINARY; PRT; 186 AA.
ID Q8KV2
AC Q8KV2;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE 60 kDa chaparonin (Fragment).
GN CPN60.
OS uncultured pig faeces bacterium.
OC Bacteria; environmental samples.
OX NCBI_TaxID=190966;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22034982; PubMed=12039767;
RA Hill J.E., Seipp R.P., Betts M., Hawkins L., Van Kessel A.G.,
RA Crosby W.L., Hemmingsen S.M.;
RT "Extensive profiling of a complex microbial community by high-
RT throughput sequencing."
RL Appl. Environ. Microbiol. 68:3055-3066(2002).
DR EMBL: AF437042; AAM49303.1; -.
DR InterPro: IPR002423; Cpn60/TCP-1.
DR Pfam: PF00118; cpn60_TCP1; 1.
KW ATP-binding; Chaparone.
FT NON TER 1
FT NON TER 186
SQ SEQUENCE 186 AA; 20151 MW; 7BE9B65B3329114D CRC64;

Query Match 60.0%; Score 6; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ITVEDA 9
Db 85 ITVEDA 90

```



## RESULT 13

Q8KH08

ID Q8KH08 PRELIMINARY; PRT; 186 AA.

AC Q8KH08;

DT 01-OCT-2002 (TREMblrel. 22, Created)

DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)

DE 60 kDa chaperonin (Fragment).

GN CPM60.

OS unclutered pig faeces bacterium.

OC Bacteria; environmental samples.

OC NCBI\_TaxID=190966;

OX [1]

RN SEQUENCE FROM N.A.

RP MEDLINE=22034982; PubMed=12039767;

RA Hill J.E., Seipp R.P., Betts M., Hawkins L., Van Kessel A.G.,

RA Crosby W.L., Hemmingsen S.M.,

RT "Extensive profiling of a complex microbial community by high-

RT throughput sequencing."

RT Appl. Environ. Microbiol. 68:3055-3066(2002).

DR EMBL; AF436897; AAA49158.1; -

DR EMBL; AF437007; AAA49258.1; -

DR EMBL; AF437129; AAA49390.1; -

DR InterPro; IPR002423; Cpm60/TCP-1.

DR Pfam; PF00118; cpm60\_TCP1; 1.

DR ATP-binding; Chaperone.

KM NON\_TER 1 1

FT NON\_TER 1 1

SQ SEQUENCE 186 AA; 20201 MW; 7BB9B5B26290CEF CRC64;

Query Match 60.0%; Score 6; DB 2; Length 186;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 4 ITVEDA 9

DB 85 ITVEDA 90

## RESULT 14

Q9PST2

ID Q9PST2 PRELIMINARY; PRT; 186 AA.

AC Q9PST2;

DT 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)

DE RNA binding/signal transduction protein Qxi-4.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus

OC NCBI\_TaxID=9031;

OX [1]

RN SEQUENCE FROM N.A.

RP STRAIN=Hubbard White Mountain;

RC MEDLINE=98208728; PubMed=9547512;

RA Mezquita J., Pau M., Mezquita C.,

RT "Four isoforms of the signal-transduction and RNA-binding protein Qxi

RT expressed during chicken spermatogenesis."

RL Mol. Reprod. Dev. 50:70-78(1998).

CC -1- SIMILARITY: CONTAINS 1 KH DOMAIN.

DR EMBL; U86767; AAD06624.1; -

DR InterPro; IPR004087; KH\_dom.

DR Pfam; PF00013; KH; 1.

DR SMART; SM00322; KH; 1.

SQ SEQUENCE 186 AA; 21303 MW; A36DEB0C164CABC1 CRC64;

Query Match 60.0%; Score 6; DB 13; Length 186;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 4 ITVEDA 9

DB 85 ITVEDA 90

DB 155 ITVEDA 160

## RESULT 15

Q9AB28

ID Q9AB28 PRELIMINARY; PRT; 187 AA.

AC Q9AB28;

DT 01-JUN-2001 (TREMblrel. 17, Created)

DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)

DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)

DE Hypothetical protein CC0406.

GN CC0406.

OS Caulobacter crescentus.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;

OC Caulobacteraceae; Caulobacter.

OC NCBI\_TaxID=155892;

OX [1]

RN SEQUENCE FROM N.A.

RP STRAIN=ATCC 19089 / CB15;

RC MEDLINE=21173698; PubMed=11259647;

RA Nieman W.C., Feldlyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,

RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,

RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,

RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,

RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,

RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,

RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,

RT "Complete genome sequence of Caulobacter crescentus."

RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).

DR EMBL; AE005713; AAK2393.1; -

DR TIGR; CC0406; -

KM Hypothetical protein; Complete proteome.

SQ SEQUENCE 187 AA; 20381 MW; F9617F6B2E2FCB85 CRC64;

Query Match 60.0%; Score 6; DB 16; Length 187;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 4 ITVEDA 9

DB 87 ITVEDA 92

Search completed: December 22, 2003, 11:57:17  
Job time : 37 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:55:16 ; Search time 22 Seconds  
(without alignments)  
19.232 Million cell updates/sec

Title: US-09-745-763-36\_COPY\_231\_240

Perfect score: 10  
Sequence: 1 TACTIVEDAE 10

Scoring table: OLIGO  
Gapop 60.0 , Gapept 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

1: Issued Patents AA:\*  
2: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/6CTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	100.0	473	4	US-09-482-273-120
2	6	60.0	351	4	US-09-252-991A-31251
3	6	60.0	371	2	US-08-837-593-8
4	6	60.0	387	2	US-08-837-593-4
5	6	60.0	393	2	US-08-837-593-3
6	6	60.0	532	2	US-08-579-777A-2
7	5	50.0	35	4	US-09-152-060-112
8	5	50.0	98	4	US-09-511-881A-14
9	5	50.0	103	6	5171684-5
10	5	50.0	115	4	US-09-252-991A-27427
11	5	50.0	129	3	US-09-011-074-2
12	5	50.0	139	4	US-09-015-189-4
13	5	50.0	145	3	US-09-011-074-6
14	5	50.0	150	3	US-09-239-909-2
15	5	50.0	158	4	US-09-663-600A-104
16	5	50.0	180	4	US-09-252-991A-28297
17	5	50.0	185	4	US-09-107-532A-6934
18	5	50.0	186	4	US-09-328-352-6232
19	5	50.0	204	3	US-09-328-352-7175
20	5	50.0	224	3	US-09-091-889-10
21	5	50.0	247	4	US-09-679-279-19
22	5	50.0	264	4	US-09-328-352-6490
23	5	50.0	282	4	US-09-328-352-8229
24	5	50.0	293	4	US-09-252-991A-22215
25	5	50.0	295	1	US-09-570-842-2
26	5	50.0	295	1	US-09-570-842-2
27	5	50.0	295	1	US-08-700-576-2

28	5	50.0	297	4	US-09-252-991A-28307	Sequence 28307, A
29	5	50.0	318	4	US-09-134-001C-4156	Sequence 4156, Ap
30	5	50.0	319	4	US-09-724-510-2	Sequence 2, Appli
31	5	50.0	319	4	US-09-723-216-2	Sequence 2, Appli
32	5	50.0	319	4	US-09-675-227-2	Sequence 2, Appli
33	5	50.0	328	2	US-08-828-242-1	Sequence 1, Appli
34	5	50.0	328	2	US-08-910-9278-1	Sequence 1, Appli
35	5	50.0	328	3	US-09-206-499-1	Sequence 1, Appli
36	5	50.0	328	3	US-09-270-370-1	Sequence 1, Appli
37	5	50.0	332	4	US-09-134-001C-4323	Sequence 4323, Ap
38	5	50.0	341	4	US-09-107-532A-7087	Sequence 7087, Ap
39	5	50.0	343	4	US-09-967-908A-8	Sequence 8, Appli
40	5	50.0	351	4	US-08-311-731A-28	Sequence 28, Appli
41	5	50.0	362	4	US-09-328-352-7264	Sequence 7264, Ap
42	5	50.0	363	4	US-09-967-908A-10	Sequence 10, Appli
43	5	50.0	373	4	US-09-967-908A-4	Sequence 4, Appli
44	5	50.0	375	2	US-08-506-864A-1	Sequence 1, Appli
45	5	50.0	375	2	US-08-851-968-1	Sequence 1, Appli

#### ALIGNMENTS

```
RESULT 1
US-09-482-273-120
; Sequence 120, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 120
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (473)
; OTHER INFORMATION: Xaa equals stop translation
US-09-482-273-120
Query Match 100.0%; Score 10; DB 4; Length 473;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 10; Conservativity 0; Mismatches 0; Indels 0;
Ox 1 TACTIVEDAE 10
Db 231 TACTIVEDAE 240
RESULT 2
US-09-252-991A-31251
; Sequence 31251, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 31251  
LENGTH: 351  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-31251

Query Match 60.0%; Score 6; DB 4; Length 351;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITVEDA 9  
Db 89 ITVEDA 94

RESULT 3  
US-08-837-593-8  
Sequence 8, Application US/08837593  
Patent No. 5977442  
GENERAL INFORMATION:  
APPLICANT: Klessig, Daniel F.  
APPLICANT: Zhang Zhunqun  
TITLE OF INVENTION: "No. 5977442el Salicylic Acid Induced  
TITLE OF INVENTION: Map Kinase and its Use for Enhanced Disease Resistance in Plant  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dann, Dorfman, Herrell and Skillman,  
ADDRESS: 1601 Market Street, Suite 720  
CITY: Philadelphia  
STATE: PA  
COUNTRY: United States of America  
ZIP: 19103-2307  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/837,593  
FILING DATE: April 21, 1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/029,805  
FILING DATE: October 25, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Reed, Janet E.  
REGISTRATION NUMBER: 36,252  
REFERENCE/DOCKET NUMBER: Rutgers 97-0016  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 563-4100  
TELEFAX: (215) 563-4044  
TELEX:  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 371 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
US-08-837-593-8  
Query Match 60.0%; Score 6; DB 2; Length 371;  
Best Local Similarity 100.0%; Pred. No. 22;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 4 ITVEDA 9  
Db 313 ITVEDA 318

RESULT 4  
US-08-837-593-4  
Sequence 4, Application US/08837593  
Patent No. 5977442  
GENERAL INFORMATION:  
APPLICANT: Klessig, Daniel F.  
APPLICANT: Zhang Zhunqun  
TITLE OF INVENTION: "No. 5977442el Salicylic Acid Induced  
TITLE OF INVENTION: Map Kinase and its Use for Enhanced Disease Resistance in Plant  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dann, Dorfman, Herrell and Skillman,  
ADDRESS: 1601 Market Street, Suite 720  
CITY: Philadelphia  
STATE: PA  
COUNTRY: United States of America  
ZIP: 19103-2307  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/837,593  
FILING DATE: April 21, 1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/029,805  
FILING DATE: October 25, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Reed, Janet E.  
REGISTRATION NUMBER: 36,252  
REFERENCE/DOCKET NUMBER: Rutgers 97-0016  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 563-4100  
TELEFAX: (215) 563-4044  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 387 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
US-08-837-593-4  
Query Match 60.0%; Score 6; DB 2; Length 387;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 4 ITVEDA 9  
Db 329 ITVEDA 334  
RESULT 5  
US-08-837-593-3  
Sequence 3, Application US/08837593  
Patent No. 5977442  
GENERAL INFORMATION:  
APPLICANT: Klessig, Daniel F.

APPLICANT: Zhang Zhugun  
TITLE OF INVENTION: "No. 5977442el Salicylic Acid Induced  
TITLE OF INVENTION: Map Kinase and its Use for Enhanced Disease Resistance in Plant  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dann, Dorfman, Herrell and Skillman,  
ADDRESSEE: P.C.  
STREET: 1601 Market Street, Suite 720  
CITY: Philadelphia  
STATE: PA  
COUNTRY: United States of America  
ZIP: 19103-2307  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/837,593  
FILING DATE: April 21, 1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/029,805  
FILING DATE: October 25, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Reed, Janet E.  
REGISTRATION NUMBER: 36,252  
REFERENCE/DOCKET NUMBER: Rutgers 97-0016  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 563-4100  
TELEFAX: (215) 563-4044  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 393 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
US-08-837-593-3

Query Match 60.0%; Score 6; DB 2; Length 393;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITVEDA 9  
Db 334 ITVEDA 339

RESULT 6  
US-08-579-777A-2  
Sequence 2, Application US/08579777A  
Patent No. 5912153  
GENERAL INFORMATION:  
APPLICANT: Enderlin, Carol S.  
APPLICANT: Seliterminkoff, Claude P.  
TITLE OF INVENTION: (1,3)-B-glucan synthase genes and  
TITLE OF INVENTION: inducible inhibition of fungal growth using the antisense  
TITLE OF INVENTION: constructs derived therefrom.  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Davis, Graham & Stubbs L.L.P.  
STREET: 370 Seventeenth Street  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.A.  
ZIP: 80201-0185  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/579,777A  
FILING DATE: 28-DEC-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/155,004  
FILING DATE: 18-NOV-1993  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 892-7370  
TELEFAX: (303) 893-1379  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 532 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Neurospora crassa  
STRAIN: wild-type 74-DR 23-1VA  
DEVELOPMENTAL STAGE: Mycelial  
US-08-579-777A-2

Query Match 60.0%; Score 6; DB 2; Length 532;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TVEDAE 10  
Db 520 TVEDAE 525

RESULT 7  
US-09-152-060-112  
Sequence 112, Application US/09152060  
Patent No. 6448230  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 28 Human Secreted Proteins  
FILE REFERENCE: P2003Pl.US  
CURRENT APPLICATION NUMBER: US/09/152,060  
EARLIER FILING DATE: 1998-09-11  
EARLIER APPLICATION NUMBER: PCT/US98/04858  
EARLIER FILING DATE: 1998-03-12  
EARLIER APPLICATION NUMBER: 60/040,762  
EARLIER FILING DATE: 1997-03-14  
EARLIER APPLICATION NUMBER: 60/040,710  
EARLIER FILING DATE: 1997-03-14  
EARLIER APPLICATION NUMBER: 60/050,934  
EARLIER FILING DATE: 1997-05-30  
EARLIER APPLICATION NUMBER: 60/048,100  
EARLIER FILING DATE: 1997-05-30  
EARLIER APPLICATION NUMBER: 60/048,357  
EARLIER FILING DATE: 1997-05-30  
EARLIER APPLICATION NUMBER: 60/048,189  
EARLIER FILING DATE: 1997-05-30  
EARLIER APPLICATION NUMBER: 60/057,765  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/068,368  
EARLIER FILING DATE: 1997-12-19  
NUMBER OF SEQ ID NOS: 118  
SOFTWARE: Patent Ver. 2.0  
SEQ ID NO 112  
LENGTH: 35  
TYPE: PRT

ORGANISM: Homo sapiens  
US-09-152-060-112

Query Match 50.0%; Score 5; DB 4; Length 35;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVED 8  
|||  
Db 6 ITVED 10

RESULT 8  
US-09-511-881A-14  
Sequence 14, Application US/09511881A  
Patent No. 6495674

GENERAL INFORMATION:  
APPLICANT: SALK INSTITUTE  
APPLICANT: LEMKE, GREGORY  
APPLICANT: NGUYEN, ANDREW  
APPLICANT: KRAPP, RALF  
TITLE OF INVENTION: ELECTINS AND THEIR USE  
FILE REFERENCE: SALIKINS.001A  
CURRENT APPLICATION NUMBER: US/09/511,881A  
CURRENT FILING DATE: 2000-02-25  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14  
LENGTH: 96  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (1) --(98)  
OTHER INFORMATION: Pleckstrin homology domain  
US-09-511-881A-14

Query Match 50.0%; Score 5; DB 4; Length 98;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVED 8  
|||  
Db 56 ITVED 60

RESULT 9  
5171684-5  
Patent No. 5171684  
APPLICANT: YEN, KWANG-MU; BLATT, LAWRENCE M.; KARL, MICHAEL R.  
TITLE OF INVENTION: BIOCONVERSIONS CATALYZED BY THE TOULUENE  
MONOOXYGENASE OF PSEUDOMONAS MENDOCINA KR-1  
NUMBER OF SEQUENCES: 41  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/590,374  
FILING DATE: 28-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 177,631  
FILING DATE: 05-APR-1988  
SEQ ID NO: 5  
LENGTH: 103  
5171684-5

Query Match 50.0%; Score 5; DB 6; Length 103;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVED 8  
|||  
Db 40 ITVED 44

RESULT 10

US-09-252-991A-27427  
Sequence 27427, Application US/09252991A  
Patent No. 6551795

GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 27427  
LENGTH: 115  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27427

Query Match 50.0%; Score 5; DB 4; Length 115;  
Best Local Similarity 100.0%; Pred. No. 93;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 VEDAE 10  
|||  
Db 109 VEDAE 113

RESULT 11  
US-09-011-074-2  
Sequence 2, Application US/09011074  
Patent No. 6133005  
GENERAL INFORMATION:  
APPLICANT: Pousetka, Annemarie  
APPLICANT: Coy, Johannes  
TITLE OF INVENTION: TRANSKETOLASE-RELATED PROTEIN  
FILE REFERENCE: 8484-0039-999  
CURRENT APPLICATION NUMBER: US/09/011,074  
CURRENT FILING DATE: 1998-05-21  
EARLIER APPLICATION NUMBER: PCT/DE96/01401  
EARLIER FILING DATE: 1996-07-26  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 129  
TYPE: PRT  
ORGANISM: Homo sapien  
US-09-011-074-2

Query Match 50.0%; Score 5; DB 3; Length 129;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ITVED 8  
|||  
Db 64 ITVED 68

RESULT 12  
US-09-015-189-4  
Sequence 4, Application US/09015189A  
Patent No. 6410263  
GENERAL INFORMATION:  
APPLICANT: Wallis, Nicola G.  
TITLE OF INVENTION: No. 6410263el Histidine Kinase  
FILE REFERENCE: GMS0008  
CURRENT APPLICATION NUMBER: US/09/015,189A  
CURRENT FILING DATE: 1998-01-29  
EARLIER APPLICATION NUMBER: 60/037,856  
EARLIER FILING DATE: 1997-02-07  
NUMBER OF SEQ ID NOS: 6

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-015-189-4
```

```
Query Match          50.0%; Score 5; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 1,1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      6 VEDAE 10
      |||||
Db      69 VEDAE 73
```

```
RESULT 13
US-09-011-074-6
; Sequence 6, Application US/09011074
; Patent No. 6133005
; GENERAL INFORMATION:
; APPLICANT: Pousetka, Annemarie
; APPLICANT: Coy, Johannes
; TITLE OF INVENTION: TRANSKETOLASE-RELATED PROTEIN
; FILE REFERENCE: 8484-0039-999
; CURRENT APPLICATION NUMBER: US/09/011,074
; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: PCT/DE96/01401
; EARLIER FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-011-074-6
```

```
Query Match          50.0%; Score 5; DB 3; Length 145;
Best Local Similarity 100.0%; Pred. No. 1,1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      4 ITVED 8
      |||||
Db      80 ITVED 84
```

```
RESULT 14
US-09-239-909-2
; Sequence 2, Application US/09239909
; Patent No. 6284952
; GENERAL INFORMATION:
; APPLICANT: Kumo Petrochemical Co. Ltd.
; TITLE OF INVENTION: Transgenic Plants with Divergent SCAM4 or SCAM5 Gene to Achieve M
; FILE REFERENCE: P99-2-6
; CURRENT APPLICATION NUMBER: US/09/239,909
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: EP 99300136.1
; EARLIER FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 2
; LENGTH: 150
; TYPE: PRT
; ORGANISM: G. max calmodulin4 (SCAM4)
US-09-239-909-2
```

```
Query Match          50.0%; Score 5; DB 3; Length 150;
Best Local Similarity 100.0%; Pred. No. 1,2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3 CITVE 7
      |||||
Db      27 CITVE 31
```

```
RESULT 15
US-09-663-600A-104
; Sequence 104, Application US/09663600A
; Patent No. 6573068
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclet, Aymeric
; APPLICANT: Bouguetel, Lydie
; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS
; FILE REFERENCE: 31.US3.CIP
; CURRENT APPLICATION NUMBER: US/09/663,600A
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/191,997
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent .pm
; SEQ ID NO 104
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -37...-1
US-09-663-600A-104
```

```
Query Match          50.0%; Score 5; DB 4; Length 158;
Best Local Similarity 100.0%; Pred. No. 1,2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 ACTIV 6
      |||||
Db      37 ACTIV 41
```

Search completed: December 22, 2003, 11:58:24  
Job time : 22 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2003, 11:56:11 ; Search time 30 Seconds  
(without alignments)  
62.253 Million cell updates/sec

Title: US-09-745-763-36\_COPY\_231\_240  
Perfect score: 10  
Sequence: 1 TACTIVEDAR 10

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 696363 seqs, 186758610 residues

Word size : 0

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubppaa/PCR\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/2/pubppaa/PCRUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubppaa/US09C\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubppaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubppaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100.0	472	9	US-09-745-763-36	Sequence 36, App1
2	100.0	473	11	US-09-984-271-120	Sequence 1200, App
3	60.0	432	9	US-09-815-242-11700	Sequence 11700, A
4	60.0	433	12	US-09-946-290-14	Sequence 14, App1
5	60.0	610	9	US-09-815-242-5559	Sequence 5559, Ap
6	60.0	618	9	US-09-815-242-12567	Sequence 12567, A
7	60.0	618	9	US-09-815-242-12970	Sequence 12970, A
8	60.0	618	9	US-09-815-242-12970	Sequence 12970, A
9	60.0	618	9	US-09-815-242-12970	Sequence 12970, A
10	60.0	618	9	US-09-815-242-12970	Sequence 12970, A
11	60.0	618	9	US-09-815-242-12970	Sequence 12970, A
12	60.0	618	9	US-09-815-242-12970	Sequence 12970, A
13	60.0	618	9	US-09-815-242-12970	Sequence 12970, A
14	60.0	618	9	US-09-815-242-12970	Sequence 12970, A
15	60.0	618	9	US-09-815-242-12970	Sequence 12970, A

16	5	50.0	35	9	US-09-852-659A-112	Sequence 112, App
17	5	50.0	35	10	US-09-852-797-112	Sequence 112, App
18	5	50.0	50	12	US-10-371-634-26	Sequence 26, App1
19	5	50.0	55	9	US-09-864-761-34650	Sequence 34650, A
20	5	50.0	60	12	US-10-029-386-29649	Sequence 29649, A
21	5	50.0	94	15	US-10-204-887-114	Sequence 114, App
22	5	50.0	104	15	US-10-177-478-38	Sequence 38, App1
23	5	50.0	104	15	US-10-156-761-8337	Sequence 8337, App
24	5	50.0	128	12	US-10-029-386-32159	Sequence 32159, A
25	5	50.0	129	10	US-09-862-658-5	Sequence 5, App1
26	5	50.0	129	15	US-10-051-767-11	Sequence 11, App1
27	5	50.0	129	15	US-10-175-696-26	Sequence 26, App1
28	5	50.0	132	10	US-09-738-626-4098	Sequence 4098, App
29	5	50.0	132	15	US-10-156-761-12474	Sequence 12474, A
30	5	50.0	150	9	US-09-925-301-1372	Sequence 1372, App
31	5	50.0	158	12	US-10-319-763-104	Sequence 104, App
32	5	50.0	176	9	US-09-780-717-2	Sequence 2, App1
33	5	50.0	184	10	US-09-984-245-218	Sequence 218, App
34	5	50.0	184	11	US-09-966-262-218	Sequence 218, App
35	5	50.0	184	11	US-09-983-966-218	Sequence 218, App
36	5	50.0	184	15	US-10-143-090-218	Sequence 218, App
37	5	50.0	190	10	US-09-738-626-4637	Sequence 4637, App
38	5	50.0	191	12	US-09-981-151A-83	Sequence 83, App1
39	5	50.0	198	10	US-09-925-300-1199	Sequence 1199, App
40	5	50.0	200	11	US-09-764-891-4187	Sequence 4187, App
41	5	50.0	210	16	US-10-080-170-58	Sequence 58, App1
42	5	50.0	217	10	US-09-738-626-5801	Sequence 5801, App
43	5	50.0	228	12	US-10-029-386-32350	Sequence 32350, A
44	5	50.0	240	15	US-10-106-696-6361	Sequence 6361, App
45	5	50.0	244	9	US-09-867-550-1200	Sequence 1200, App

#### ALIGNMENTS

RESULT 1  
US-09-745-763-36  
Sequence 36, Application US/09745763  
Patent No. US20020065394A1

#### GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth  
McCoy, John M.  
LaValle, Edward R.  
Collins-Racie, Lisa A.  
Evans, Cheryl  
Merberg, David  
Treacy, Maurice  
Spaulding, Vikki  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM  
NUMBER OF SEQUENCES: 219  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140

#### COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/745,763  
FILING DATE: 18-Jun-2000  
CLASSIFICATION: <Unknown>

#### ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.  
REGISTRATION NUMBER: 41,323  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8284  
TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 472 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 36:  
US-09-745-763-36

Query Match 100.0%; Score 10; DB 9; Length 472;  
Best Local Similarity 100.0%; Pred. No. 0.0035;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TACTIVEDAE 10  
|||||  
Db 231 TACTIVEDAE 240

RESULT 2  
US-09-984-271-120  
Sequence 120, Application US/09984271  
Publication No. US20030040088A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 71 Human Secreted Proteins  
FILE REFERENCE: P2030P1  
CURRENT APPLICATION NUMBER: US/09/984,271  
CURRENT FILING DATE: 2001-10-29  
PRIOR APPLICATION NUMBER: 09/482,273  
PRIOR FILING DATE: 2000-01-13  
PRIOR APPLICATION NUMBER: PCT/US99/15849  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: 60/092,921  
PRIOR FILING DATE: 1998-07-15  
PRIOR APPLICATION NUMBER: 60/092,922  
PRIOR FILING DATE: 1998-07-15  
PRIOR APPLICATION NUMBER: 60/092,956  
PRIOR FILING DATE: 1998-07-15  
NUMBER OF SEQ ID NOS: 267  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 120  
LENGTH: 473  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (473)  
OTHER INFORMATION: Xaa equals stop translation  
US-09-984-271-120

Query Match 100.0%; Score 10; DB 11; Length 473;  
Best Local Similarity 100.0%; Pred. No. 0.0035;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TACTIVEDAE 10  
|||||  
Db 231 TACTIVEDAE 240

RESULT 3  
US-09-815-242-11700  
Sequence 11700, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA 011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11700  
LENGTH: 432  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-815-242-11700

Query Match 60.0%; Score 6; DB 9; Length 432;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITVEDA 9  
|||||  
Db 89 ITVEDA 94

RESULT 4  
US-09-946-290-14  
Sequence 14, Application US/09946290  
Publication No. US2003021475A1  
GENERAL INFORMATION:  
APPLICANT: Roberts, Christopher J.  
TITLE OF INVENTION: METHODS FOR IDENTIFYING PATHWAY-SPECIFIC REPORTERS AND  
FILE REFERENCE: 9301-040  
CURRENT APPLICATION NUMBER: US/09/946,290  
CURRENT FILING DATE: 2001-09-05  
PRIOR APPLICATION NUMBER: US/09/282,243  
PRIOR FILING DATE: 1999-03-31  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 14  
LENGTH: 433  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-946-290-14

Query Match 60.0%; Score 6; DB 12; Length 433;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ITVEDA 9  
|||||  
Db 307 ITVEDA 312

RESULT 5  
US-09-815-242-5559  
Sequence 5559, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.

```

; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5559
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5559
```

```

Query Match      60.0%; Score 6; DB 9; Length 610;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      5 TVEDAE 10
Db      595 TVEDAE 600
```

```

RESULT 6
US-09-815-242-12567
; Sequence 12567, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5559
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5559
```

```

; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12567
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12567
```

```

Query Match      60.0%; Score 6; DB 9; Length 618;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      5 TVEDAE 10
Db      603 TVEDAE 608
```

```

RESULT 7
US-09-815-242-12970
; Sequence 12970, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12970
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12970
```

```

Query Match      60.0%; Score 6; DB 9; Length 618;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      5 TVEDAE 10
Db      603 TVEDAE 608
```

```

RESULT 8
US-09-771-161A-163
; Sequence 163, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
```

;; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
;; FILE REFERENCE: 802620-2005.1  
;; CURRENT APPLICATION NUMBER: US/09/771,161A  
;; CURRENT FILING DATE: 2001-01-26  
;; PRIOR APPLICATION NUMBER: 09/724,676  
;; PRIOR FILING DATE: 2000-11-28  
;; PRIOR APPLICATION NUMBER: 136776  
;; PRIOR FILING DATE: 2000-06-15  
;; PRIOR APPLICATION NUMBER: 135619  
;; PRIOR FILING DATE: 2000-04-12  
;; NUMBER OF SEQ ID NOS: 273  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 163  
;; LENGTH: 842  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-771-161A-163

Query Match 60.0%; Score 6; DB 10; Length 842;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 TVEDAE 10  
|||||  
Db 70 TVEDAE 75

RESULT 9  
US-10-005-983-2  
;; Sequence 2, Application US/10005983  
;; Publication No. US20020116730A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Allen, Keith D.  
;; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING PERK PROTEIN  
;; FILE REFERENCE: R-517  
;; CURRENT APPLICATION NUMBER: US/10/005,983  
;; CURRENT FILING DATE: 2001-11-07  
;; PRIOR APPLICATION NUMBER: US 60/246,676  
;; PRIOR FILING DATE: 2000-11-07  
;; PRIOR APPLICATION NUMBER: US 60/311,018  
;; PRIOR FILING DATE: 2001-08-08  
;; PRIOR APPLICATION NUMBER: US 60/324,765  
;; PRIOR FILING DATE: 2001-09-24  
;; PRIOR APPLICATION NUMBER: US 60/326,148  
;; PRIOR FILING DATE: 2001-09-28  
;; NUMBER OF SEQ ID NOS: 4  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 2  
;; LENGTH: 1114  
;; TYPE: PRT  
;; ORGANISM: Mus musculus  
US-10-005-983-2

Query Match 60.0%; Score 6; DB 14; Length 1114;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 TVEDAE 10  
|||||  
Db 67 TVEDAE 72

RESULT 10  
US-09-771-161A-253  
;; Sequence 253, Application US/09771161A  
;; Patent No. US20020110811A1  
;; GENERAL INFORMATION:  
;; APPLICANT: LEVINE, et al.  
;; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
;; FILE REFERENCE: 802620-2005.1  
;; CURRENT APPLICATION NUMBER: US/09/771,161A  
;; CURRENT FILING DATE: 2001-01-26

;; PRIOR APPLICATION NUMBER: 09/724,676  
;; PRIOR FILING DATE: 2000-11-28  
;; PRIOR APPLICATION NUMBER: 136776  
;; PRIOR FILING DATE: 2000-06-15  
;; PRIOR APPLICATION NUMBER: 135619  
;; PRIOR FILING DATE: 2000-04-12  
;; NUMBER OF SEQ ID NOS: 273  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 253  
;; LENGTH: 1115  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-771-161A-253

Query Match 60.0%; Score 6; DB 10; Length 1115;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 TVEDAE 10  
|||||  
Db 70 TVEDAE 75

RESULT 11  
US-10-296-540-2  
;; Sequence 2, Application US/10296540  
;; Publication No. US20030215827A1  
;; GENERAL INFORMATION:  
;; APPLICANT: JULIER, C, cile  
;; APPLICANT: DELEPINE, Marc  
;; APPLICANT: NICOLINO, Marc  
;; TITLE OF INVENTION: MUTATED EUKARIOTIC TRANSLATION INITIATION FACTOR 2 ALPHA KINASE 3,  
;; TITLE OF INVENTION: MULTIPLE EPIPHYSEAL DYSPLASIA (WOLCOTT-RALLISON SYNDROME)  
;; FILE REFERENCE: 344 061 - US  
;; CURRENT APPLICATION NUMBER: US/10/296,540  
;; CURRENT FILING DATE: 2002-11-25  
;; PRIOR APPLICATION NUMBER: EP 00/401 436  
;; PRIOR FILING DATE: 2000-05-23  
;; PRIOR APPLICATION NUMBER: EP 00/402 707  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: PCT/IB 01/01 153  
;; PRIOR FILING DATE: 2001-05-23  
;; NUMBER OF SEQ ID NOS: 105  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 2  
;; LENGTH: 1115  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-296-540-2

Query Match 60.0%; Score 6; DB 12; Length 1115;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 TVEDAE 10  
|||||  
Db 70 TVEDAE 75

RESULT 12  
US-10-350-405-201  
;; Sequence 201, Application US/10350405  
;; Publication No. US20030215894A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Niman, Henry L.  
;; TITLE OF INVENTION: Polypeptide-induced Monoclonal Receptors  
;; TITLE OF INVENTION: to Protein ligands  
;; FILE REFERENCE: TSRI 35.5 CON 7/LIG  
;; CURRENT APPLICATION NUMBER: US/10/350,405  
;; CURRENT FILING DATE: 2003-01-24  
;; PRIOR APPLICATION NUMBER: 09/427,576  
;; PRIOR FILING DATE: 1999-10-26

```

; PRIOR APPLICATION NUMBER: 08/461,563
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: 08/294,879
; PRIOR FILING DATE: 1994-08-23
; PRIOR APPLICATION NUMBER: 08/054,864
; PRIOR FILING DATE: 1993-04-28
; PRIOR APPLICATION NUMBER: 07/900,502
; PRIOR FILING DATE: 1992-06-16
; PRIOR APPLICATION NUMBER: 07/780,415
; PRIOR FILING DATE: 1991-10-22
; NUMBER OF SEQ ID NOS: 227
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201
; LENGTH: 12
; TYPE: PRT
; ORGANISM: retrovirus
US-10-350-405-201

Query Match          50.0%; Score 5; DB 12; Length 12;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 CITE 7
        |||||
Db       1 CITE 5

RESULT 13
US-10-057-789-239
; Sequence 239, Application US/10057789
; Publication No. US20030082522A1
; GENERAL INFORMATION:
; APPLICANT: Paul Haynes
; APPLICANT: Jing Wei
; APPLICANT: John Yates
; TITLE OF INVENTION: DIFFERENTIAL LABELING FOR QUANTITATIVE
; TITLE OF INVENTION: ANALYSIS OF COMPLEX PROTEIN MIXTURES
; FILE REFERENCE: NADII.022A
; CURRENT APPLICATION NUMBER: US/10/057,789
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 60/264,576
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/305,232
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 239
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 4
; OTHER INFORMATION: Xaa = Modified Cysteine
US-10-057-789-239

Query Match          50.0%; Score 5; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 VEDA 10
        |||||
Db       11 VEDA 15

RESULT 14
US-10-212-628-239
; Sequence 239, Application US/10212628
; Publication No. US20030087329A1
; GENERAL INFORMATION:
; APPLICANT: Paul Haynes
; APPLICANT: Jing Wei
```

```

; APPLICANT: John Yates
; APPLICANT: Nancy Andon
; TITLE OF INVENTION: DIFFERENTIAL LABELING FOR QUANTITATIVE
; TITLE OF INVENTION: ANALYSIS OF COMPLEX PROTEIN MIXTURES
; FILE REFERENCE: NADII.022CPI
; CURRENT APPLICATION NUMBER: US/10/212,628
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 60/264,576
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/305,232
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 10/057,789
; PRIOR FILING DATE: 2002-01-25
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 239
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 4
; OTHER INFORMATION: Xaa = Modified Cysteine
US-10-212-628-239

Query Match          50.0%; Score 5; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 VEDA 10
        |||||
Db       11 VEDA 15

RESULT 15
US-09-853-161-112
; Sequence 112, Application US/09853161
; Patent No. US20020076756A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003p3
; CURRENT APPLICATION NUMBER: US/09/853,161
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 112
; LENGTH: 35
; TYPE: PRT
```

ORGANISM: Homo sapiens  
US-09-853-161-112

Query Match 50.0%; Score 5; DB 9; Length 35;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 ITVED 8  
|||  
Db 6 ITVED 10

Search completed: December 22, 2003, 11:59:08  
Job time : 31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: December 22, 2003, 20:22:58 ; Search time 4346 Seconds  
(without alignments)  
4443.015 Million cell updates/sec

Title: US-09-745-763-36  
Sequence: 1 MKFLIFAFGCVHLLSLCSG.....AVNAVSVYVADMEMLPRS 472

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+g2n.model -DEV=xlp  
-Q=/cgn2\_1/USPRO.spool/p/US09745763/runat.22122003.113317.25251/app.query.fasta.1.647  
-DB=GenBml -QPMT=faetap -SUFRTX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09745763 @CGN 1.1.3745 @runat.22122003.113317.25251 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenBml.\*  
1: gb\_da.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_cm.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_rts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_da.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_rts.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_hcg\_hum.\*  
31: em\_hcg\_inv.\*  
32: em\_hcg\_other.\*  
33: em\_hcg\_mus.\*  
34: em\_hcg\_pin.\*  
35: em\_hcg\_rtd.\*  
36: em\_hcg\_mam.\*  
37: em\_hcg\_vrt.\*  
38: em\_sy.\*  
39: em\_hugo\_hum.\*  
40: em\_hugo\_mus.\*  
41: em\_hugo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2417	100.0	1472	9	AF107834	AF107834 Homo sapi
2	2417	100.0	1851	6	BD106411	BD106411 Secreted
3	2417	100.0	1860	6	AX006440	AX006440 Sequence
4	2417	100.0	1928	9	BC020689	BC020689 Homo sapi
5	2410	99.7	1778	6	BD105816	BD105816 Secretory
6	2402	99.4	1860	6	BD127520	BD127520 Primer fo
7	2402	99.4	1860	9	AK075132	AK075132 Homo sapi
8	2384	98.6	1794	9	AF119386	AF119386 Homo sapi
9	2133	88.2	1596	10	AF107835	AF107835 Mus muscu
10	2133	88.2	1796	10	BC037067	BC037067 Mus muscu
11	2117	87.6	1726	10	AF131077	AF131077 Rattus no
12	2116	87.5	1778	6	AX400829	AX400829 Sequence
13	2116	87.5	1778	10	AF097723	AF097723 Rattus no
14	2029	83.9	1716	10	AF009513	AF009513 Mus muscu
15	870	36.0	609	6	BD059610	BD059610 Secreted
16	771	31.9	642	6	BD125210	BD125210 Primer fo
17	771	31.9	642	6	BD126476	BD126476 Primer fo
18	769.5	31.8	1729	3	AF077194	AF077194 Acanthoch
19	745	30.8	156702	2	AC084223	AC084223 Homo sapi
20	745	30.8	157927	9	AP006278	AP006278 Homo sapi
21	745	30.8	171301	9	AC010859	AC010859 Homo sapi
22	701	29.0	423	9	AF107833	AF107833 Homo sapi
23	662	27.4	10792	1	AE015790	AE015790 Shevaneli
24	640	26.5	10487	1	AE005922	AE005922 Caulobact
25	623	25.8	417	10	AF107832	AF107832 Mus muscu
26	606	25.1	167883	2	AC121026	AC121026 Rattus no
27	606	25.1	218874	2	AC110103	AC110103 Rattus no
28	600	24.8	214287	2	AC129951	AC129951 Mus muscu
29	513	21.2	10927	1	AE012431	AE012431 Xanthomon
30	499.5	20.7	11712	1	AE011976	AE011976 Xanthomon
31	491	20.3	462	6	BD105811	BD105811 Secretory
32	491	20.3	462	6	E62849	E62849 Secretary p
33	479	19.8	314	6	BD058492	BD058492 Secreted
34	393	16.3	144703	9	AP003112	AP003112 Homo sapi
35	351	14.5	317	6	BD076934	BD076934 5' EST of
36	346	14.3	152176	9	AP003117	AP003117 Homo sapi
37	346	14.3	171035	9	AP003111	AP003111 Homo sapi
38	340	14.1	218874	2	AC110103	AC110103 Rattus no
39	334	13.8	214287	2	AC129951	AC129951 Mus muscu
40	332	13.7	203469	2	AC123654	AC123654 Mus muscu
41	302	12.5	228034	2	AC123681	AC123681 Mus muscu
42	298	12.3	166050	2	AC013817	AC013817 Homo sapi
43	298	12.3	177008	2	AC011175	AC011175 Homo sapi
44	297	12.3	240178	2	AC094173	AC094173 Rattus no
45	297	12.3	288136	2	AC120489	AC120489 Rattus no

RESULT 1

## ALIGNMENTS

AF107834 1472 bp mRNA linear PRI 12-JUL-1999  
 LOCUS Homo sapiens clone LCH1 aminopeptidase mRNA, complete cds.  
 DEFINITION AF107834  
 ACCESSION AF107834  
 VERSION AF107834.1 GI:5442029  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 1472)  
 AUTHORS Liu, C.H., Lin, B.Y. and Chang, L.Y.  
 TITLE Cloning of the human aminopeptidase gene  
 JOURNAL unpublished  
 REFERENCE 2 (bases 1 to 1472)  
 AUTHORS Liu, C.H., Lin, B.Y. and Chang, L.Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia  
 Sinica, RM 433, 128, Yen-Chun-Yuan Road SEC 2, Taipei 11529, Taiwan  
 FEATURES  
 source  
 1. 1472  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="LCH1"  
 /tissue\_type="liver"  
 /dev\_stage="fetus"  
 54..1472  
 /note="secretory protein"  
 /codon\_start=1  
 /product="aminopeptidase"  
 /protein\_id="AAD43214.1"  
 /db\_xref="GI:5442030"  
 /translation="MKELIFAPFGVNHLSGSKATKNGISKRTSEIKESIASCS  
 DVAATILNLAAYGKAQNSRYERLALVDVPRLSGSNLEKAIQIMONQOGLK  
 VHLSEPRIPHEWREGESAVMLEPRHKTAIILGSSISCTPEGIITAEVLVTSDELQ  
 RRASERARKIVVNOPYINRGTVQFQGAVEAKVGLASLRSVASFISPHNG  
 IOERODGVPRKIPACITVEDAMSRMSHGKIYIOLKMAKTPPDTSFNTVAEIT  
 GSKYDEQVVLVSGHLDSDVNGGANDDGGAPISWEALSILKIDGLRKRRLVLT  
 AEKGGVGAFOYQLHKVNIISYSLVMSDGTFLPTGLQGTGSKARALIEEWSLL  
 QPLNTQVLSHGEQDINFWIQAGVPGASLLDLDLYKFFPHSHGDTWMDPKOMV  
 AAAYWAVVSYSVADEEMLEPRPS"  
 BASE COUNT 411 a 325 c 381 g 355 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 8,836-176 Length: 1472  
 Score: 2417.00 Matches: 472  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-745-763-36 (1-472) x AF107834 (1-1472)  
 Oy 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCySerGly 20  
 Db 54 ATGAATATCTTATCTTGCATTTTCCGTGCTGTTCACTTTTATCCCTGCTCTG 113  
 Oy 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluIle 40  
 Db 114 AAAGCTATATGCAAGATGGCATCTCTAAGAGCATTTTGAAGAATAAAGAAATA 173  
 Oy 41 AlaSerCyGlyAspValAlaLysAlaIleLeuLeuAlaValIleGlyLysAlaGln 60  
 Db 174 GCCAGCTGTGGAGATGTTGCTTAAGCAATCATCACTAGCTGTTATGTTAAAGCCAG 233  
 Oy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
 Db 234 AACGATCTATAGACGATGGCATCTTGGTGTGACTGTGTGGACCAAGACTGAGTGC 293  
 Oy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100  
 |||||||

Db 294 TCCAAGAACCTAGAAAAAGCATCAATTATGTATCCAAAACTGCAGCAAGATGGCTG 353  
 Oy 101 GluLysValHisLeuGluProValArgIleProHisIleProGluArgGlyGluIleSerAla 120  
 Db 354 GAGAAAGTCACTCGAGGAGGATGAGATACCCCATCTGGAGAGGAGGAGAAATCAAGCT 413  
 Oy 121 ValMetLeuGluProValGlnIleLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
 Db 414 CTGATGCTGAGACCAAAATTCATATAGATAGCCATCTCTGGGTCTGGACAGCATTTGGG 473  
 Oy 141 ThrProProGluGlyIleThrAlaGluValLeuValIleThrSerPheAspGluLeuGln 160  
 Db 474 ACTCTCTCAGAGGACATTACAGCAGAAAGTTCTGTGTGTGACCTCTTTCATGTAAGCTCAG 533  
 Oy 161 ArgArgAlaSerGluAlaArgGlyValIleValValIleThrAsnGlnProTyrIleAsnTyr 180  
 Db 534 AGAAGGCTCTCAGAGCAAGAGGAGAAATGTTGTTATTAACCAACTTATCATCAATCACTAC 593  
 Oy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaAlaLysValGlyAlaLeu 200  
 Db 594 TCAGAGACGGTGCATACCAAGCCAGGGGGCGGTGGAAGCTGCCAAGTTGGGGCTTTG 653  
 Oy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisIleGlnIleGln 220  
 Db 654 GCATCTCTCATTCGATCCGTGGCTCTCTCTCCATCTACAGTCCCTCACAGGCAATTCAG 713  
 Oy 221 GluLysThrIleAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
 Db 714 GAATACCAAGATGGCGGCCCAAGATTCCAAAGCTGTATTCAGGTGAAGACGACGAA 773  
 Oy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
 Db 774 ATGATGTCAGAAATGGCTTCTCATAGGATCAAAATTCATTCATCACTAAAGATGGGGCA 833  
 Oy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280  
 Db 834 AAGACCTACCAAGATCTGATTCCTTCAACACTGTACAGAGATCACTGGGAGCAAAATAT 893  
 Oy 281 ProGluGluValValIleLeuValIleSerGlyHisLeuAspSerTyrAspValGlyGlnAla 300  
 Db 894 CCAAAACAGGTGTACTGTCAGTGCAGTGCATCTGCAGCTGGAGTGTGGACGGGTGCC 953  
 Oy 301 MetAspAspGlyGlyGlyAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320  
 Db 954 ATGATATATGGCGGTGAGACCTTATATCATGGAAACATCTCATTATTAAGATCTT 1013  
 Oy 321 GlyLeuArgProLysArgThrLeuArgLeuValIleuThrAlaGluGluGlnGly 340  
 Db 1014 GGGCTGCTGCCAAAGAGGACTCTGGCTGTGCTGTGCACTGCAGAAACAAAGGTGGA 1073  
 Oy 341 ValGlyAlaPheGlnTyrTyrGlnIleHisLysValAsnIleSerAsnTyrSerLeuVal 360  
 Db 1074 GTTGTCCTTCACGATATACGTATACAGTACACAGGATAAATTTTCAACTACAGTCTG 1133  
 Oy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluLys 380  
 Db 1134 ATGAGTCTGACGAGGAGAACCTTCTTACCACTGGCTGCATTCATCGCAGAGGAAAG 1193  
 Oy 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400  
 Db 1194 GCCAGGCGCATATGAGAGAGGATATAGCTGTGCAGGCCCTCCATATATCACTCAGGTC 1253  
 Oy 401 LeuSerHisGlyGluGlyThrAspIleAsnPheThrIleGlnAlaGlyValProGlyAla 420  
 Db 1254 CTGAGCATGAGAGAGGAGACAGATCACTTTTGATCCAAAGCTGAGTGCCTGGAGCC 1313  
 Oy 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisIleSerHisGlyAspThrMet 440  
 Db 1314 AGTCTACTGATGATCTTATACAGATATTTCTTCTTCATCACTCCACGAGACACCATG 1373  
 Oy 441 ThrValMetAspProLysGlnMetAsnValAlaAlaIleValTPAlaValIleSerTyr 460  
 Db 1374 ACTGTATGATCCAAAGCAGATGATGTTGCTGCTCTGTGGCTGTGGCTGTTCTTAT 1433  
 |||||||



Qy 461 ValValAlaAspMetGluMetLeuProArgSer 472  
Db 1434 GTTGTTCAGACATGGAAGAAATGCTGCTAGTCC 1469

RESULT 2  
BD106411 1851 bp DNA linear PAT 18-SEP-2002  
LOCUS Secreted proteins and polynucleotides encoding them.  
DEFINITION BD106411  
ACCESSION BD106411 GI:23201229  
VERSION UP 2002503955-A/2.  
KEYWORDS Chlamydia sp.  
SOURCE Chlamydia sp.  
ORGANISM Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
REFERENCE 1 (bases 1 to 1851)  
AUTHORS Jacobo, K., McCoy, J.M., Lavallie, E.R., Racie, L.A., Merberg, D., Treacy, M., Spaulding, V., and Agostino, M.J.  
TITLE Secreted proteins and polynucleotides encoding them  
JOURNAL Patent: JP 2002503955-A 2 05-FEB-2002;  
GENETICS INSTITUTE INC  
COMMENT PN JP 2002503955-A/2  
PD 05-FEB-2002  
PR 20-MAR-1998 JP 1998545874  
PR 20-MAR-1997 US 08/822167, 19-MAR-1998 US 09/044466 PI  
KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI  
DAVID MERBERG,  
PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC  
C12N15/12, C07K14/47, A61K8/17  
CC Strandedness: Double;  
Topology: Linear;  
FH Key Location/Qualifiers.

FEATURES  
source 1..1851  
/organism="Chlamydia sp."  
/mol\_type="genomic DNA"  
/db\_xref="taxon:35827"

BASE COUNT 531 a 413 c 438 g 469 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1,18e-175 Length: 1851  
Score: 2417.00 Matches: 472  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-09-745-763-36 (1-472) x BD106411 (1-1851)

Qy 1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGly 20  
Db 99 ATGAAATTCCTTATCTTCGATCTTTCGCTGCTGCTTATCCCTGCTGCTGG 158

Qy 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40  
Db 159 AAACCTATATCCAGATGCGATCTCTRAAGGACCTTTGAAGAAATTAAGAAATA 218

Qy 41 AlaSerCyGlyAspValAlaIleLeuLeuAlaValIleLysAlaGln 60  
Db 219 GCCAGCTGTGGAGATGTTGCTAAAGCATCATCACTGCTTTATGTTAAAGCCAG 278

Qy 61 AsnArgSerTyrGluValGluLeuValAspThrValGlyProArgLeuSerGly 80  
Db 279 AACGATCTCAATGAGCGATGCTGCTGTTGATCTGTGAGCCAGACTGAGTGGC 338

Qy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100  
Db 339 TCCAGAACTTCAAGAAAGCCATCCCAATTTATGCCAAACCTGCGAGCAAGATGGCTG 398

Qy 101 GluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluGluSerIle 120  
Db 399 GAGAAAGTTCACTGAGCCAGTGAATATCCCACTGCGAGAGGGGAGAGAAATCAAGCT 458

Qy 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
Db 459 GTGATGTGAGCGCAAGAAATTCATTAAGATAGCATCTGGGTCTTGCGACGACATGGG 518

Qy 141 ThrProGluGluIleThrAlaGluValLeuValIleHisSerPheAspGluLeuGln 160  
Db 519 ACTCTCCAGAAAGCATTTACAGCAAGATGCTGGTGGTGAACCTTTGCATGAATCGAG 578

Qy 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAspTyr 180  
Db 579 AGAAGGCTCCAGAAAGCAAGAGGAAGATTGTTTATTAACCAACCTTACATCACTAC 638

Qy 181 SerArgThrValGlnTyrArgTyrGlnGlyAlaValAlaGluAlaLysValGlyAlaLeu 200  
Db 639 TCAGAGCGGTGCAATTCAGAACGACGAGGCGCGTGGAGCTCCAGAGTGGGGCTTTG 658

Qy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisTyrGlyIleGln 220  
Db 699 GCATCTCTCATTCGATCCGTGGCTCTCTTCATCTACAGTCTCTCAACAGGATTCAG 758

Qy 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
Db 759 GAATACCGAGATGCGCGCCCAAAATTCACACGCTGTATTACGGTGAAGATGCAGAA 818

Qy 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
Db 819 ATGATGTCAGAAATGCTTCTTCATGGGATCAAAATTCATTCAGCTTAAGATGGGGCA 878

Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280  
Db 879 AAGCCTACCCAGATCTGATTCCTTCAACACTGTAGCAAGATCATCTGGAGCAATAT 938

Qy 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTyrAspValGlyGlnGlyAla 300  
Db 939 CCAAGACAGGTGTACGTGCAAGTGCAGTGCAGTGCAGTGCAGTGTGGGACGGGGCC 998

Qy 301 MetAspAspGlyGlyGlyAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320  
Db 999 ATGATATGCGCGTGGAGCTTTTATTCATGGGACACTCTCATTTAATAAGATCTT 1058

Qy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTyrThrAlaGluGluGlnGly 340  
Db 1059 GGGGTGCTCCAAAGAGACTGCGCGCTGCTGCTGCACTGAGAGAAACAAGTGA 1118

Qy 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360  
Db 1119 GTTGTGCTTCCAGATTTATTCAGTTACACAGGTAAATTTTCCAACTACAGTCTGCTG 1178

Qy 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380  
Db 1179 ATGAGTCTTACCGAGGAACTTTTACCACTGGGCTGAATTCATCTGCGAGGAAAG 1238

Qy 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuProLeuAsnIleThrGlnVal 400  
Db 1239 GCCAGGCGCATATGAGAGAGGATTATAGCTGCTGCGACCTCATATCACTCAAGTC 1298

Qy 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTyrIleGlnAlaGlyValProGlyAla 420  
Db 1299 CTGAGCCATGAGAGAGGACAGACATCACTTTTGGATCCCAAGTGGAGTGGCGAGCC 1358

Qy 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440  
Db 1359 AGTCTACTTATATGACTTATACAGATATTTCTTCCATCACTCCACGAGAGACCATG 1418

Qy 441 ThrValMetAspProLysGlnMetAsnValAlaAlaIleValTyrAlaValIleSerTyr 460  
Db 1419 ACTGTCAATGATCCAAAGCAGATGAATGCTGCTGCTGCTTTGGCGCTGTGTTCTTAT 1478

Qy 461 ValValAlaAspMetGluGluMetLeuProArgSer 472  
Db 1479 GTTGTTCAGACATGGAAGAAATGCTGCTAGTCC 1514

RESULT 3  
AX006440  
LOCUS AX006440 1860 bp DNA linear PAT 06-SEP-2000  
DEFINITION Sequence 5 from Patent WO0004157.  
ACCESSION AX006440  
VERSION AX006440.1 GI:9994575  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
1 Neefs, J.M., Peeters, D.C. and Pangalos, M.  
AUTHORS Cloning and characterization of novel mammalian peptidases  
TITLE Patent: WO 0004157-A 5 27-JAN-2000;  
JANSEN PHARMACEUTICA NV (BE); NEEFS JEAN MARC EDMOND FERNAND (BE);  
PEETERS DANIELLE CELINE GEORGE (BE); PANGALOS MENILAS (GB)  
FEATURES  
source  
1..1860  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 518 a 420 c 451 g 471 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,196-175 length: 1860  
Score: 2417.00 Matches: 472  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
US-09-745-763-36 (1-472) x AX006440 (1-1860)  
QY 1 MetLysPheLeuIlePheAlaPheheGlyGlyValHisLeuLeuSerLeuCySerGly 20  
DB 125 ATGAATTCCTTATCTTCGATTTTTCGATGTTTCACTTTTATCCCTGCTGCTGG 184  
QY 21 LysAlaIleCybLysAsnGlyIleSerLysArgThrPheGluGluIleLeuGluIle 40  
DB 185 AAAGCTAATATGCAAAATGGCATCTCTTAAGAGACTTTTGAAGAAATAAAGAAATA 244  
QY 41 AlaSerCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60  
DB 245 GCCAGCTGTGAGATGTTGCTTAAGCATCATCACTGCTTTATGTTAAAGCCAG 304  
QY 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
DB 305 AACAGATCTTAAGACGATGGCACTTCGTTGATATGTTGAGACCAAGCTAGTGGC 364  
QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100  
DB 365 TCCAGAACTTAAGAAAGCCATCCAAATTATGATCCAAACCTCCAGACAGATGGCTG 424  
QY 101 GluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluGluSerAla 120  
DB 425 CAGAAAGTTCACTGAGACCACTGAGATACCCCACTGGAGAGGGAGAAAGATCACT 484  
QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
DB 485 GTGATGCTGAGACCAAGATTCATTAAGATGACCATCTGGGCTTGGCAGCAGCTGGG 544  
QY 141 ThrProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160  
DB 545 ACTCTCTCAAGAGGATTAACAGCAAGTTCGTGGTGAGACCTCTTTCGATGAATGAG 604  
QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAspTyr 180  
DB 605 AGAAGGCTCTCAAGAGCAAGAGGAGATGTTGTTTAAACCAACTTACACTACACTAC 664  
QY 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaValValValGlyValLeu 200

DB 665 TCAAGACGGTGCAATACCAAGCAGAGGGCGGTGAGAGCTGCCAAGTTGGGGCTTTG 724  
QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220  
DB 725 GCATCTCTCATTCATCCGTGGCTCTCTTCTCCATTCACAGTCTCCACACAGGATTCAG 784  
QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
DB 785 GAATACCAAGATGGCGTCCCAAGATTCCAAGACCTGTTATACGTTGAGGAAGATGCAGAA 844  
QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
DB 845 ATGATGTCAGAAAGGCTTCTCATGGATCAAAATTCATTCAGCTCAAGAAAGATGGGGCA 904  
QY 261 LysThrTyrProAspThrArgSerPheAspThrValAlaGluIleThrGlySerLysTyr 280  
DB 905 AAGACCTACCAAGATCTGATCTTCAACACGTGTACAGAGATCATCTGGAGCAAAATAT 964  
QY 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300  
DB 965 CCAAGACAGGTGTACTGCTGCTCAGTGACATCTGCACAGCTGGATGTTGGCAGAGTGGC 1024  
QY 301 MetAspAspGlyGlyGlyAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320  
DB 1025 ATGATGATGCGCGGTGAGGCTTTATATCATGGAGACACTCTCATTTAAGATCTT 1084  
QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrPheAlaGluGluGlnGly 340  
DB 1085 GGGCTGGCTCCAAAGAGAGACTCTGGCTGCTGTCTGAGCTGCAGAGAAACAGAGTGA 1144  
QY 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360  
DB 1145 GTTGCTGCTTCCAGATTTATCATGATTACAGATTAATTTCCAACTACAGCTCTGCTG 1204  
QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380  
DB 1205 ATGAGAGCTGACGAGAGAACCTTCTTACCACTGGCTGCATTCATCTGCAGTGAAGA 1264  
QY 381 AlaArgAlaIleMetGluGluValMetSerLeuGlnProLeuAsnIleThrGlnVal 400  
DB 1265 GCCAGGCGCATCATGAGAGAGTTATAGCTGTGAGGCCCTCAATATCATCTCAGGTC 1324  
QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420  
DB 1325 CTGAGCATGAGAGAGGACAGACATCACTTTTGATCCAGCTGAGTGGCCGAGGCC 1384  
QY 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440  
DB 1385 AGTCTACTGATGACTTATACAGTATTTCTTCTTCATCATCACTCCACGAGACACCATG 1444  
QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaValAlaValValValSerTyr 460  
DB 1445 ACTGTCAATGATCCAAAGACAGATGATGTTGCTGCTGTGTGGCTGTGTTCTTAT 1504  
QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472  
DB 1505 GTTCTTCAGACATGAGAAAGATGCTGCTAGTCC 1540  
RESULT 4  
BC020689 1928 bp mRNA linear PRI 22-JAN-2002  
LOCUS BC020689  
DEFINITION Homo sapiens, plasma glutamate carboxypeptidase, clone MGC:22418  
IMAGE:4251802, mRNA, complete cds.  
ACCESSION BC020689  
VERSION BC020689.1 GI:18088383  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
1 Strausberg, R.  
TITLE Direct Submission

## JOURNAL

Submitted (03-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC/DCTD/DTF  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mdickson@stanford.edu](mailto:mdickson@stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 35 Row: f Column: 13  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7706386.

## FEATURES

## source

Location/Qualifiers  
1..1928  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/db\_xref="MIM:425180"  
/clone="MGC:22418 IMAGE:4251802"  
/issue\_type="Skin, melanotic melanoma, high MDR."  
/clone\_id="NIH\_MGC\_62"  
/lab\_host="DH10B"  
/note="vector: pDNR-LIB"  
167..1585  
/codon\_start=1  
/product="plasma glutamate carboxypeptidase"  
/protein\_id="AAH20689.1"  
/db\_xref="GI:18088384"  
/translation="MKFLIFAPFGVHLISCSGKAIKNGISKRTPEIKETIASCS  
DVAKRIINLAVYKQNSRYRLALNDVTTPRSGSNLEKAIQIQWQNGLEK  
VHLEPVRIPMERGESAVMLEPRHKAIIIGLSISITPEEGITAEVLTSTFDELQ  
RRSEARGLIVNQPINYSRTVQYRQGAIVEAKVGLASLRSVASFISYSPHTG  
IOEYODGVKIPITACITVEDAEEMMSRMSHGIKIVIQKMAKTYPDIDSNFTVAIT  
GSKYDEQVIVSGHLSMDVGQANDDGGAFISWEALSLKDLGLRKRRLRLVLT  
AEQGGVGAFOYLOLHKYNISVSLVMSDAGTLPGLIOETSGEKARAIEMEMSLI  
QPLNITQYLSHGEGTDINFIQAGVPGASLDDLYKTRFFHSHSGDTVTWDPKQWV  
AAAVWAVSYVVADEMLPRS"

## CDS

## BASE COUNT

554 a 435 c 466 g 473 t

## ORIGIN

Alignment Scores:  
Pred. No.: 1.24e-175 Length: 1928  
Score: 2417.00 Matches: 472  
Percent Similarity: 100.00% Conservatave: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-745-763-36 (1-472) x BC020689 (1-1928)

Qy 1 MetLysPheLeuLeuPheAlaPhePheGlyGlyValHisLeuLeuSerLeuCySerGly 20  
Db 167 ATGAAATTCCTTATCTTCGATTTTCGTCGTTCACCTTTATCCCTGCTCTGGG 226  
Qy 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40  
Db 227 AAAGCTATATCAAGAAATGCGATCTCTAAGAGCACTTTTGAAGAAATATAAGAAATA 286  
Qy 41 AlaSerCyGlyAspValAlaLysAlaIleLeuLeuAlaValAlaTyrGlyValAlaGln 60  
Db 287 GCCAGCTGTGGAGATGTTGCTTAAGCAATCACTTACCTGTTTAAAGTAAAGCCAG 346  
Qy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80

Db 347 AACAGATCTATGAGCGATTCGACCTCTGCTGATGATCTGTGAGCCCAACTGAGTGGC 406  
Qy 81 SerLysAsnLeuGlyLysAlaIleGlnIleMetTyrGlnLysLeuGlnLysGlyLeu 100  
Db 407 TCCAAAGAACCTAGAAAAGCCCAATTAATATGACCAAAACCTGCGACAGATGGGCTG 466  
Qy 101 GlnLysValHisLeuGluProValArgIleProHisTrpGluArgGlyGluGlnSerAla 120  
Db 467 GAGAAATTCACCTGAGGCCAGTGAATATCCCACTGGAGAGAGAGAGAAATTCAGCT 526  
Qy 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
Db 527 GTGATGCTGGAGCCAAAGATTCATTAATGATGATGATGATGATGATGATGATGATGATG 586  
Qy 141 ThrProGluGluGlyIleThrAlaGluValLeuValValHisSerPheAspGluLeuGln 160  
Db 587 ACTCTCCAGAAAGCATTTACAGCAAGAGTTCTGTGTGACCTCTTTCGATGAACTGAG 646  
Qy 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180  
Db 647 AGAAGGCTCCAGAAAGCAAGCAAGGAGAGATTTGTTTATTAACCACTTACATCACTAC 706  
Qy 181 SerArgThrValGlnTyrArgThrGlnGlyAlaValGluAlaLysValGlyAlaLeu 200  
Db 707 TCAAGACGGTGCAATACCGAAGCAAGGAGGCGGTGAGAGTCCAGATTGAGGCTTTG 766  
Qy 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisTrpGlyIleGln 220  
Db 767 GCATCTCTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 826  
Qy 221 GlnTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
Db 827 GAATACCAAGATGGCGGCCCAAGATTCACAGACCTGTATTACGCTGGAAGATCGAGA 886  
Qy 241 MetLysSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
Db 887 ATGATGTCAGAAATGGCTTCATGATGATCAAAATGTGATTCATTCATTCATTCATTCAT 946  
Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280  
Db 947 AAGGCTTACCAAGATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1006  
Qy 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300  
Db 1007 CCAGAACGGTTGATCGTGTAGTGAACATCTGACAGCTGGAGTGTGGCAGAGGCTCC 1066  
Qy 301 MetAspAspGlyGlyAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320  
Db 1067 ATGATGATGCGCGGTGAGGCTTTATATCATGGGAACACTCTCATTTAAGATCTT 1126  
Qy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrAlaGluGluGlnGlyGly 340  
Db 1127 GGGCTGCTGCCAAAGAGAGACTCTGCGCTGTCTGTGACTGCAGAAAGAGAGTGGGA 1186  
Qy 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValLeuIleSerAsnTyrSerLeuVal 360  
Db 1187 GTTGCTGCTTCAGATATACATTAACATTAACAGGATTAATATTTCAACTACAGTCTG 1246  
Qy 361 MetGluSerAspAlaGlyTyrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380  
Db 1247 ATGAGCTGTAGCGAGAGAACTTTATACCACTGGCTGCATTCATCGAGAGGAAAG 1306  
Qy 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400  
Db 1307 GCCAGGCGCATCAGGAGGAGGATTAAGCTGCTGCAAGCCCTCAATATCACTCAGGTC 1366  
Qy 401 LeuSerHisGlyGluGlyTyrAspIleAsnPheThrIleGlnAlaGlyValProGlyAla 420  
Db 1367 CTGAGCCATGAGAGGAGGAGACATCACTTTGATTCCAAGCTGAGGCTGAGGCC 1426  
Qy 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440

DB	1427	AGCTCACTTGATGACTTATACAGATTTCTTCTTCATCACTCCAGGAGACACATG	1488
OY	441	ThrValMetAspProCysGlnMetAsnValAlaAlaAlaValTrpAlaValIserTyr	460
Db	1487	ACTGCATGATGATCCAAAGACAGATGATGTGTCTGCTGTTGGCTGTGTCTTAT	1546
OY	461	ValValAlaAspMetCysGlnMetLeuProArgSer	472
Db	1547	GTTGTTCAGACATGAAAGAAATGCTGCTAGTCC	1582
RESULT 5	BD105816		
LOCUS	BD105816	1778 bp	DNA linear PAT 18-SEP-2002
DEFINITION	Secretory protein.		
ACCESSION	BD105816		
VERSION	BD105816.1 GI:23200634		
KEYWORDS	JP 2002502234-A/7.		
SOURCE	synthetic construct		
ORGANISM	artificial construct		
REFERENCE	1 (bases 1 to 1778)		
AUTHORS	Jacobs, K., McCoy, J.M., Racie, L.A., Lavallie, E.R., Merberg, D. and Spaulding, V.		
TITLE	Secretory protein		
JOURNAL	Patent: JP 2002502234-A 7 22-JAN-2002;		
COMMENT	GENETICS INSTITUTE INC PN JP 2002502234-A/7 PD 22-JAN-2002 PF 16-APR-1997 JP 1997537384 PR 18-APR-1996 US 08/534325, 13-JAN-1997 US 08/793520 PI KENNETH JACOBS, JOHN M MCCOY, LISA A RACIE, EDWARD R LAVALLIE, PI DAVID MERBERG, PI VIKKI SPAULDING PC C12N15/12, C07K14/47, A61K38/17 CC Strandedness: Double; CC Topology: linear; FH Key Location/Qualifiers.		
FEATURES	source	1..1778	Location/Qualifiers
		/organism="synthetic construct"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:32630"	
BASE COUNT	514 a 386 c 422 g 456 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	3,84e-175	Length:	1778
Score:	2410.00	Matches:	471
Percent Similarity:	99.79%	Conservative:	0
Best Local Similarity:	99.79%	Mismatches:	1
Query Match:	99.71%	Indels:	0
DB:	6	Gaps:	0
OS-09-745-763-36 (1-472) x BD105816 (1-1778)			
OY	1	MethylsPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCysSerGly	20
Db	17	ATGAATTCCTTATCTTCGCATTTTTCGTGTGTTCACCTTTATCCCTGTGCTCGG	76
OY	21	LySaIaIleCYbLYbAsnGlyIleSerLybArgThrPheGlnGlnIleLybGlnGlnIle	40
Db	77	AAAGCTATATGCAGAAATGCGATCTCTPAAAGAGACTTTTGAAGAAATAAAGAAATA	136
OY	41	AlaSerCYbGlyAspValAlaLybAlaIleIleAsnLeuAlaValTYrGlybLYbAlaGln	60
Db	137	GCCAGCTGTGGAGATGTGCTPAAACCAATCATCAACCTAGCTTTATGTATAGCCAG	196
OY	61	AsnArgSerTYrGlnArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly	80
Db	197	AACAGATCTTATGACCGATGGCACTTCTGTGTATACTGTTGGACCCAGACTAGTGC	256
OY	81	SerLybAsnLeuGlnLybAlaIleGlnIleMetTYrGlnAsnLeuGlnAspGlyLeu	100

Db	257	TCCAAGAACCTCAGAAAAAGCATCCAAATTATGTATCCAAAACCTGCAGCAAGATGGCGT	316
Oy	101	GLuLyVaHisLseuGluProValArgIleProHisTPGuaArgGlyGluGluSerAla	120
Db	317	GAGAAAGTTCACCTGGAGCCAGTGAAGATATCCCACTGGAGAGAGGAGAAAGATTCAGCT	376
Oy	121	ValMetLeuGluProAlaGlyIleHisIleValIleAlaIleLeuGlyLeuGlySerSerIleGly	140
Db	377	GTGATGCTGGAGCCMAAATTCAATMAATAGCCATCTGGGTCTTGGCAGCAGCATTTGGG	436
Oy	141	ThrProProGluGlyIleThrAlaGluValIleuValValThrSerPheAspGluLeuGln	160
Db	437	ACTCTCTCAGAAAGGCATTACAGCAGAAAGTTCTGGTGTGACCTTTTCGATGAACCTGCAG	496
Oy	161	ArgArgAlaSerGluAlaArgGlyLysIleValValTyrranginProTyrrIleAsnTyrr	180
Db	497	AGAAAGGCGCTCAGAAAGCAGAAAGGAGGAGATTTGTTTATTAACAACCTTACATCACTAC	556
Oy	181	SerArgThrValGlnTyrrArgThrGlnGlyAlaValGluAlaAlaLysValGlyAlaLeu	200
Db	557	TCAAAGCAGCGGTCAATACCGAAACCCAGCGGGCGGTGAAAGCTGCAGAGGTGGGGCTTGG	616
Oy	201	AlaSerLeuIleArgSerValAlaSerPheSerIleTyrrSerProHisThrGlyIleGln	220
Db	617	GCATCTCTCATTCATCCGTGGCTCTCTCTCCATCTACACATCTCTCAACAGGATTTACG	676
Oy	221	GluTyrrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu	240
Db	677	GAATACCAAGATGGCGGTGCCCAAAATTCCAACAGCCTGTATTAAGGTGGAAAGATGCAGAA	736
Oy	241	MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla	260
Db	737	ATGATGTCAAGAAAGGCTTCTCATGGGATCAAAATTCATTCAGCTCAAGATGGGGCA	796
Oy	261	LysThrTyrrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIleTyrr	280
Db	797	AAGACCTACCCAGATACCTGATCTCTTAACAACGTAGAGAGATCACTGGGAGCAAAATAT	856
Oy	281	ProGluGlnValValIleuValSerGlyHisIleuAspSerTPaAspValGlyGlnGlyAla	300
Db	857	CCAGAACAGGTGTACGTGCTCAGTGAACATCTGGACAGCTGGATGTTGGCCAGAGTGGCC	916
Oy	301	MetAspAspGlyGlyGlyAlaPheIleSerTPGluAlaLeuSerLeuIleLysAspLeu	320
Db	917	ATGATGATATGGCGGTGAGCCTTATATATCATGGAAACACTCTCACTTATTAAGATCTTT	976
Oy	321	GlyLeuAspArgProLysArgThrLeuArgLeuValLeuThrThrAlaGluGluGlnGlyGly	340
Db	977	GGGCTGGTCCAAAGAGAGACTGGCGGTGGTGGTGGAGTGGAGAGAACCAAGTGGCA	1036
Oy	341	ValGlyAlaPheGlnTyrrTyrrGlnLeuHisIleLysValAsnIleSerAsnTyrrSerLeuVal	360
Db	1037	GTTCGTGCTTCCAGTATATACATTACAGTTACACAGGTAATATTTCCAACTACAGCTGGTG	1096
Oy	361	MetGluSerAspAlaGlyTyrrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys	380
Db	1097	ATGAGAGTCAACGACAGAACTTCTTACCCACTGGCTGCATTCACATCGCAGTGAAGAAAG	1156
Oy	381	AlaArgAlaIleMetGluGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal	400
Db	1157	GCCAGGGCCATCATGAGAGAGGTATATAGCCTGTGCAGCCCTCCAAATATACACAGTGC	1216
Oy	401	LeuSerHisGlyGlyGluGlyTyrrAspIleAsnPheTyrrIleGlnAlaGlyValProGlyAla	420
Db	1217	CTGAGCCATGAGAAAGGAGACAGACATCACTTTTGATTCCAAGCTGGAGTGCCTGGAGCC	1276
Oy	421	SerLeuLeuAspAspLeuTyrrLysTyrrPheBebHeHisSerHisGlyAspThrMet	440
Db	1277	AGTCTACTGTATGACTTATACAAAGTATTTCTTCTTCATCACTCTCCACGGAGACACCATG	1336
Oy	441	ThrValMetAspProLysGlnMetAsnValAlaAlaAlaValTPAlaValAlaValSerTyrr	460
Db	1337	ACTGTCATGTATCCAAAGCAGATCAAAAGTTGGCTGCTCTGTTTGCGGTGTGTGTTCTTAT	1396

QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472  
 DB 1297 GTTGTTCAGACATGAGAAATGCTGCTAGTCTC 1432  
 RESULT 6  
 BDI27520 1860 bp DNA linear PAT 18-SEP-2002  
 LOCUS Primer for synthesizing full-length cDNA and use thereof.  
 DEFINITION BDI27520  
 ACCESSION BDI27520.1 GI:23222465  
 VERSION JP 2002017375-A/2951.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1860)  
 Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,  
 Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and  
 Koga,H.  
 TITLE Primer for synthesizing full-length cDNA and use thereof  
 JOURNAL Patent: JP 2002017375-A 2951 22-JAN-2002;  
 COMMENT HELIX RESEARCH INSTITUTE  
 OS Homo sapiens (human)  
 PN JP 2002017375-A/2951  
 PD 22-JAN-2002  
 PF 07-JUL-2000 JP 2000253172  
 PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO  
 PI ISHII,  
 PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI  
 SHINICHI KOJIMA,  
 PI TETSUJI OTSUKI, HISASHI KOGA  
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
 10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
 PC Primer for synthesizing full-length cDNA and use thereof FH key  
 FEATURES  
 FT CDS Location/Qualifiers  
 1..1860  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 515 a 423 c 451 g 471 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.66e-174 Length: 1860  
 Score: 2402.00 Matches: 469  
 Percent Similarity: 99.79% Conserved: 2  
 Best Local Similarity: 99.36% Mismatches: 1  
 Query Match: 99.38% Indels: 0  
 Gaps: 0  
 US-09-745-763-36 (1-472) x BDI27520 (1-1860)  
 QY 1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCysSerGly 20  
 DB 129 ATGAATTCCTATCTTCGCAATTTTCGTCGTCGTCACCTTTATCCCTGCTGCTGG 188  
 QY 21 LysAlaIleCysLysAsnGlyIleSerLysArgTrpPheGluGluIleGlyGluGlu 40  
 DB 189 AAAGCTAATATGCAAAATGGCATCTCTTAAGAGACTTTTGAAGAAATGAAGAAATA 248  
 QY 41 AlaSerCysGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60  
 DB 249 GCCAGCTGTGAGAGTGTCTTAAGCATCACTAGCTGTTATGCTTAAGCCAG 308  
 QY 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
 DB 309 AACAGATCTTATGACGATGGACCTCTGCTGTTACTGTTGAGACCAAGACTGAGTGC 368

QY 81 SerLysAsnLeuGluValAlaIleGlnIleMetTyrGlnAsnLeuGlnAspGlyLeu 100  
 DB 369 TCCAGAACCTTGAAGAAAGCATTCATATTAATGACAAACCTGCAGCAAGATGGCTG 428  
 QY 101 GluLysValHisLeuGluProValArgIleProHisTrpGluArgGlyGluSerAla 120  
 DB 429 GAGAAATTCACCTGGAGCCAGTGAATATCCCACTGGGAGGAGGAGGAGATTCAGCT 488  
 QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
 DB 489 GTGATGCTGAGCCAAAGAAATTCATTAAGATAGCCATCTCTGGGCTTGGCAGCAGATTGG 548  
 QY 141 ThrProProGluGlyIleThrAlaGluValLeuValAlaThrSerPheAspGluLeuGln 160  
 DB 549 ACTCTCCAGAGACCATTTACAGCAAGATTCGTGTGTGATGACCTTTTCATGAACTCAG 608  
 QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180  
 DB 609 AGAAGGCTCCAGAAAGCAAGAGGAGGAAATGTTGTTTATTAACCACTTACATCACTAC 668  
 QY 181 SerArgThrValGlnTyrArgTrpGlnGlyAlaValGluAlaAlaLysValGlyAlaLeu 200  
 DB 669 TCAAGGACGGTGCATATCCGAAACGAGGGCGGTGAGACTGCCAAGTTGGGGCTTTG 728  
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220  
 DB 729 GCATCTCTCATTCGATTCGGTGGCTCTTCTCCATCTACAGCTCTTCACACAGGATTCAG 788  
 QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
 DB 789 GAATACCAAGATGCGTGGCCCAAGATTCCAACACCTGTATTACGGTGAAGATGCAGAA 848  
 QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
 DB 849 ATGATGTCAAAGAAAGGCTTTCATGATGATCAAAATTCATTCAGCTTAAGATGGGCA 908  
 QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280  
 DB 909 AAGACCTACCCAGATCTGATCTCTTCAACACTGATACAGATGATCTGGAGCAAAATAT 968  
 QY 301 MetAspAspGlyGlyGlyAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320  
 DB 1029 ATGATGATGCGCGGTGAGCCTTTATATCATGGAGACACTCTCATTTAATGAATCTT 1088  
 QY 321 GlyLeuArgProLysArgTrpLeuArgLeuValLeuTrpThrAlaGluGluGlnGly 340  
 DB 1089 GGGCTGCTCCAAAGAGAGACTCTCGGCTGTGCTCTGACTGCAGAAAGACAGGTGGA 1148  
 QY 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360  
 DB 1149 GTTGCTGCTCCAGATATATCACTTACACAGGTAATATTTCCACTACAGCTGCTGGT 1208  
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380  
 DB 1209 ATGAGAGCTGACGAGAGAACTTCTTACCACTGGGCTGCAGAAATTCACCTGGCAGGA 1268  
 QY 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlyVal 400  
 DB 1269 GCCAGGGCCATCGTGAAGAGTATAGCTGCTGCAGCCCTCCATATATCTCAGGCTC 1328  
 QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420  
 DB 1329 CTGAGCCATGAGAGAGGAGACATCATCACTTTTGATTCACAGCTGGAGTGCCTGAGCC 1388  
 QY 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisHisSerHisGlyAspThrMet 440  
 DB 1389 AGTCTACTTGATGATTAATACAGATATTTCTTCTTCATCATCTCCACGAGAGACCATG 1448  
 QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaValAlaValValValSerTyr 460

Db 1449 ACTGTCATGATCCAAAGACATGATGTTGCTGCTGCTGTTGGCTGTTCTTAT 1508  
Qy 461 ValValAlaAspMetGluGluMetLeuProArgSer 472  
Db 1509 GTTGTTCAGACATGAGAAATGCTGCTAGTCC 1544

RESULT 7  
AK075132

LOCUS  
DEFINITION  
AK075132 1860 bp mRNA linear PRI 03-SEP-2002  
Homo sapiens cDNA FLJ90651 fis, clone PLACE1004482, moderately  
similar to Rattus norvegicus hematopoietic lineage switch 2 related  
protein (Hs18-1p) mRNA.

ACCESSION  
AK075132  
VERSION  
AK075132.1 GI:22761022  
KEYWORDS  
oligo-capping; fis (full insert sequence).

SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens

REFERENCE  
1 Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,  
Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,  
Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,  
Kojima, S., Nagahara, K., Masuko, Y., Ono, T., Okano, K., Yoshikawa, Y.,  
Aotsubu, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and  
Ninomiya, K.  
NEO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 1860)  
Isogai, T. and Otsuki, T.  
Direct Submission  
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,  
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@helix.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology; cDNA library construction:  
Institute of Medical Science, University of Tokyo; Laboratory of  
Genome Structure, Human Genome Center; cDNA 5'-3' end one pass  
sequencing and clone selection: Helix Research Institute (supported  
by Japan Key Technology Center etc.).  
Location/Qualifiers  
1. 1860  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="PLACE1004482"  
/tissue\_type="Placenta"  
/clone\_lib="PLACE1"  
/note="Cloning vector: pME18FLJ3"  
129. 1547  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAC11423.1"  
/db\_xref="GI:22761023"  
/translation="MKFLIPAPFGVHLISCSKATICKNGISKRTPEIRBIAISCS  
DVAKIINLAYGKQNSRYERIALVVTTPRSLSGSKLKAIDIMQNCQDGLSC  
VLEPRIPMERGESAVMLEPRTHKIAIGLSSISTPEGITAELVLTSDLELQ  
RRASARKIVVNPQYINYSRTVQRQGAVEAKVLAASLRSVAFLSYSPHTG  
IOEVODGPKIPACTIVEDAEEMSRMSHGIKIVIQMGAKTYPDTSNFTVAEIT  
GSKYDEQVVLVSGHLSWDVCGAMDGGAFFISWEALSLKDLGRKRLRLVLT  
AESGGVGAFOYVQLHKYINISYLSWMSDAGTFLPTGLQDTGSEKRAIYEEWSSL  
OPNITQVLSHGEGTDINPFIQAGVFGASLDDLYKTFPHSHGDTWTWDPKQNV  
AAAWAVSYVADMEMLPRS"

BASE COUNT 515 a 423 c 451 g 471 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.66e-174 Length: 1860  
Score: 2402.00 Matches: 469  
Percent Similarity: 99.79% Conservative: 2  
Best local Similarity: 99.36% Mismatches: 1

Query Match: 99.38% Indels: 0  
DB: 9 Gaps: 0  
US-09-745-763-36 (1-472) x AK075132 (1-1860)

Qy 1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCysSerGly 20  
Db 129 ATGAATTCCTTATCTTCGCACTTTTTCGGTGGTGTTCACCTTTATCTCTGCTGGG 188  
Qy 21 LysAlaIleCysLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGluIle 40  
Db 189 AAACCTTATATGCAAGATGGCATCTCTMAAGAGACTTTTGAAGAAATGAAGAAATA 248  
Qy 41 AlaserCysGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60  
Db 249 GCCACGCTGTGAGAGATGTTGCTTAAAGCAATCATCACTTACTGTTTATGTTAAAGCCAG 308  
Qy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
Db 309 AACGATCTTATAGAGCATGGATTGGCACTTCTGTTGATCTGTTGGACCCAGACTGAGTGGC 368  
Qy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100  
Db 369 TCCAAAGAACCTTAAAGAAAGCCATCCAAATTATGACCAAAACCTGCAGAAAGATG33CTG 428  
Qy 101 GluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluGluSerAla 120  
Db 429 GAGAAATTCACTCGAGGCCAGTGAAGATACCCCACTGGAGAGAGGAGAGAAATCACTGCT 488  
Qy 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGluLysSerSerIleGly 140  
Db 489 GTGATCTGAGAGCCAAATTCATTAAGATAGCATCTCTGGGCTTTCAGACAGCATTTGGG 548  
Qy 141 ThrProGluGluGlyIleThrAlaGluValLeuValHisSerPheAspGluLeuGln 160  
Db 549 ACTCTTCAGAAAGCATTTAAGCAAGATCTGTGTGTGATCTTTCAGATGAATCTGACG 608  
Qy 161 ArgArgLaserGluAlaArgGlyLysIleValValTyrAsnGlnProTyrIleAsnTyr 180  
Db 609 AGAAGGCTCTCAAGAGCAAGAAAGGAGGATGTTGTTTATACCACTTATCATCACTAC 668  
Qy 181 SerArgThrValGlnTyrArgThrGlnGlyValAlaGluAlaLysValGlyAlaLeu 200  
Db 669 TCAAGACGCTGCAATCCAGACGAGGCGGCTGGAACCTGCAAGTGGGCTTTG 728  
Qy 201 AlaserLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220  
Db 729 GCATCTCTCATTCGATCGTGGCTCTCTCTCTCATCTACAGTCTTCACAGAGATTACG 788  
Qy 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
Db 789 GAATACAGATGGCGCGGCCCAAGATTCCAAGCGCTGTTACGGTGAAGATGCAGAA 848  
Qy 241 MetLysSerArgPheAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
Db 849 ATGATGTCAGAAATGCGCTTCTCATGGGATCAAAATGTTCATTCAGCTTAAAGATGGGCA 908  
Qy 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280  
Db 909 AAGGCTTACCAAGATTCGATCTCTTCACACGTCGACAGATCACTGGAGCAAAATAT 968  
Qy 281 ProGluGlnValValLeuValSerGlyHisLeuAspSerTyrAspValGlyGlnGlyAla 300  
Db 969 CCAAGAACGTTGTACTGCTGACATGAGACATCTGACAGCTGGAGATGTGGCGAGGCTGCC 1028  
Qy 301 MetAspAspGlyGlyGlyAlaPheIleSerTyrGluAlaLeuSerLeuIleLysAspLeu 320  
Db 1029 ATGATGATGCGCGGTGAGCCTTATATCATAGGAAGCACTTCACCTTATTAAGATCTT 1088  
Qy 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTyrThrAlaGluGluGlnGlyGly 340  
Db 1089 GGGCTGCTGTCAAAGAGACTCTGCGGCTGTCTCTGAGCTGCAGAGAAAGAAAGGTGGA 1148

Qy		341	ValGlyAlaPheGlnTyrTyrGlnLeuHisValAsnIleSerAsnTyrSerLeuVal	360
Db		1149	GTGGTGCTTCCAGTATTATCAGTTACAAAGTAATATTTCCACTACAGTCTGGT	1208
Qy		361	MetGluSerAspAlaGlyThrPheLeuProThrglyLeuGlnPheThrglySerGluIys	380
Db		1209	ATGAGAGTCTGACGCGAGAAACCTTTCTTACCACTGGGCTGCATTCACCTGCGAGTAAG	1268
Qy		381	AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal	400
Db		1269	GCCAGGGGCATCGGAGAGAGATTATGAGCTGCGACGCCCTCAATATCATCAGGTC	1328
Qy		401	LeuSerHisGlyGluGlyThrAspIleAsnThrPrlIleGlnAlaGlyValProGlyAla	420
Db		1329	CTGAGCCATGGAGAAAGGACACATCACTTTGGATCCAAAGCTGAGAGTCCCTGGAGCC	1388
Qy		421	SerLeuLeuAspAspLeuTyrTyrIleTyrPhePhePheHisIleSerHisGlyAspThrMet	440
Db		1389	AGTCTACTGTGATGACTTATACAGATTTCTTCTTCATCATCTCCACGAGACACCATG	1448
Qy		441	ThrValMetAspProLysGlnMetCysValAlaAlaAlaValTTPAlaValAlaSerTyr	460
Db		1449	ACTGTCATGAGATCCAAAGCAGATGATTTCTGCTGCTGTTGGCTGTTGTTCTTAT	1508
Qy		461	ValValAlaAspMetGluGlnMetLeuProIleSer	472
Db		1509	GTGTTCGACATGAGAAAGATGCTGCTTGAAGTTC	1544
RESULT 8				
LOCUS	AF119386	1794 bp	mRNA	linear PRI 20-MAY-1999
DEFINITION	Homo sapiens blood plasma glutamate carboxypeptidase precursor			
ACCESSION	AF119386			
VERSION	AF119386			
KEYWORDS	AF119386.1 GI:4877697			
SOURCE				
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 1794)			
AUTHORS	Gingras, R., Richard, C., El-Alfy, M., Morales, C.R., Potier, M. and			
	Pfelezhetsky, A.V.			
	Purification, cDNA cloning, and expression of a new human blood			
	plasma glutamate carboxypeptidase homologous to			
	N-acetyl-aspartyl-alpha-glutamate			
	carboxypeptidase/prostate-specific membrane antigen			
	J. Biol. Chem. 274 (17), 11742-11750 (1999)			
JOURNAL	99223495			
MEDLINE	10206990			
PUBMED	2 (bases 1 to 1794)			
REFERENCE	Gingras, R., Richard, C., El-Alfy, M., Morales, C.R., Potier, M. and			
AUTHORS	Pfelezhetsky, A.V.			
TITLE	Direct Submission			
JOURNAL	Submitted (12-JAN-1999) Medical Genetics, Sainte-Justine Hospital,			
	Montreal University, 3175 Cote Sainte-Catherine, Montreal, QU H3T			
	1C5, Canada			
FEATURES				
source	Location/Qualifiers			
	1..1794			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/chromosome="8"			
	/map="8q22.2; between DBS257 and DBS270"			
	1..1794			
	/gene="PGCP"			
	25..1650			
	/gene="PGCP"			
	/EC_number="3.4.17.21"			
	/function="endopeptidase"			
	/note="member of M28 peptidase family; similar to			
	glutamate carboxypeptidase II and to			
	N-acetyl-aspartyl-alpha-glutamate carboxypeptidase			

	US-09-745-763-36 (1-472) x AF119386 (1-1794)	
QY	1 MetIysPheIeuIlePheAlaPhePheGlyGlyValHisIleuLeuSerIeuCysSergly 20	
Db	25 ATGAATTCCTTATCTTGCGCATTTTCGCGTGTCTACCTTTTATCCCTGTGCTCGG 84	
QY	21 LysAlaIleCysAluYsaIngIyleSerIysArqThrPheGluGluIleYsgIuGluIle 40	
Db	85 AAGCTATATGCAAGATGGCATCTCTAAGAGACTTTGAAAGAAATTAAGAAAGATA 144	
QY	41 AlISerCygeGlyAspValAlaYsaAlaIleIleAsnLeuAlaValIlyrGlyIyPsaIaGln 60	
Db	145 GCCAGGTGGAGATGTGCTAAACCAATCTCAACCACTGTTTATGTGTAAGCCAG 204	
QY	61 AsnArGSerTyGluArGLeuAlaLeuLeuValAspThrValGlyProArGLeuSergly 80	
Db	205 AACAGATCTTAAGAGCATGGCATCTTGCTGTTACTGTGTGAACCCAGACTGAGTGGC 264	
QY	81 SerIysAsnLeuGluIySaAlaIleGlnIleMetTyrGlnAsnLeuGlnIngInaSpGlyLeu 100	
Db	265 TCCAGAGACCTGAAAAAAGCCATCCAAATTTATGTAACCAAACTGCAGCAAGATGGGCTG 324	
QY	101 GluIyValHisIleuGluProValArgIleProHisStrpGluArGgIyGluGluSerAla 120	



```

Db      325 GAGAAAGTTCACTTGAGCCAGTGAATACCCCACTGGGAGAGGAGAAATCAGCT 384
Qy      121 ValMetLeuGluProArgIleHisIleValIleLeuGlySerSerIleGly 140
Db      385 GTGAGCTGGAGCCAGAAATTCATAGATAGACATCTGGGCTTGGGACGACGCTTGGG 444
Qy      141 ThrProGluGlyIleThrAlaGluValIleValIleThrSerPheAspIleLeuGln 160
Db      445 ACTCTCCAGAAAGCATTAACAGCAGAGTTCGTGGTGGACCTCTTGGATGAATGACG 504
Qy      161 ArgAlaGluSerGluAlaArgGlyIleValIleValIleThrGlnProTyrIleAsnTyr 180
Db      505 AGAAGGCGCTCAGAAAGCAGAGGAAAGATTTGTTTATACCAACTTACCTACATCAC 564
Qy      181 SerArgThrValGlnTyrArgThrGlnGlyAlaValIleAlaIleValGlyAlaLeu 200
Db      565 TCAGGAGCGTGCAATACCGAAGCGAGGCGGCGGAGGAGCGCCAGGCGGCGCTTGG 624
Qy      201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisThrGlyIleGln 220
Db      625 GCATCTCTCATTCGATCCGTCGCTCCTTCTCCATCTACAGTCTCCACAGCATTTGAG 684
Qy      221 GluTyrGlnAspGlyValProIleIleProThrAlaCysIleThrValGluAspAlaGlu 240
Db      685 GAATACCAAGATGCGCGTCCCAAGATTCACACGCTGATTAACGCGAAGATGACAGAA 744
Qy      241 MetMetSerArgMetAlaSerHisGlyIleIleValIleGlnLeuIleMetCysIleVal 260
Db      745 ATGATGTCAGAAATGCTTCTCATGAGCAAAATGTCATTCACCTAAAGTGGGCGCA 804
Qy      261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerIleTyr 280
Db      805 AAGACCTACCCAGATACGATGATCTTCAACACTAGACAGATGACGCGGAGCAAAAT 864
Qy      281 ProGluGlnValIleValIleSerGlyIleLeuAspSerTyrPaspValIleGlnIleVal 300
Db      865 CCAGAACAGTGTGTCTGCTGAGTCTGACATCTGACAGCTGGAGTGGGACGCGTCC 924
Qy      301 MetAspAspGlyIleValIleAlaPheIleSerTyrGluAlaLeuSerLeuIleIleValAspLeu 320
Db      925 ATGATGATGAGCGGCGGAGCTTATATCATGAGGAGCACTCTCACTTAATTAAGATCTT 984
Qy      321 GlyLeuArgProIleArgThrLeuArgLeuValIleTyrThrAlaGluGlnIleGly 340
Db      985 GGGCGCGCTCCAAAGAGATCTGCGGCTGCTGCTGAGACGACGAGAAAGAGTGGGA 1044
Qy      341 ValGlyAlaPheGlnTyrTyrGlnLeuHisIleValIleSerAsnTyrSerLeuVal 360
Db      1045 GTTGGTGCCTTCCAGTATATATGATTAACACAGTAAATATTTCCAACTACAGTCTGGTG 1104
Qy      361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGluIle 380
Db      1105 ATGAGTCTGAGGAGGAACTTCTTACCCACTGGGCTGCAATTACCTGGGACGTGAAAG 1164
Qy      381 AlaAlaGalaIleMetGluGluValIleMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400
Db      1165 GCCAGGCGCATATGAGAGAGGTTATGACCTGCTGACGCCCTCAATATCATCAAGTGC 1224
Qy      401 LeuSerHisGlyGlnIleThrAspIleAsnPheTyrIleGlnAlaGlyValProGlyAla 420
Db      1225 CTGAGCCCATGAGAGGAGCAGACATCAACTTTGGATCCAACTGAGAGTGGCTGGAACC 1284
Qy      421 SerLeuLeuAspAspLeuTyrTyrTyrPhePhePheHisIleSerHisGlyAspThrMet 440
Db      1285 AGTCTACTGATGATTAACAAAGATTTCTTCTTCATCACTCCACAGGAGACACATG 1344
Qy      441 ThrValMetAspProIleGlnMetAsnValAlaAlaIleValIleValIleValIleSerTyr 460
Db      1345 ACTGTCATGATCCAG--CAGATGATGATGCTGCTGCTGTTGGGCTGTTCTTAT 1403
Qy      461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db      1404 GTTGTTCAGACATGAGAAATGCTGCTAGTCTC 1439

```

```

RESULT 9
AF107835
LOCUS      1596 bp mRNA linear ROD 12-JUL-1999
DEFINITION Mus musculus clone LCH-M1 aminopeptidase mRNA, complete cds.
ACCESSION AF107835
VERSION    AF107835.1 GI:5442031
KEYWORDS
SOURCE
ORANISM    Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 1596)
  Liu, C.H., Lin, B.Y. and Chang, L.Y.
  Cloning of the mouse aminopeptidase gene
  Unpublished
  2 (bases 1 to 1596)
  Liu, C.H., Lin, B.Y. and Chang, L.Y.
  Direct Submission
  Submitted (19-NOV-1998) Institute of Biomedical Sciences, Academia
  Sinica, RM 433, 128 Yen-Chiun-Yuan Road SEC 2, Taipei 11529, Taiwan
FEATURES
  source
  1..1596
    /organism="Mus musculus"
    /mol_type="mRNA"
    /db_xref="taxon:10090"
    /clone="LCH-M1"
    /tissue_type="liver"
    /note="secretory protein"
    /product="aminopeptidase"
    /codon_start=1
    /protein_id="FAD63215.1"
    /db_xref="GI:5442032"
  /translation="MRSFLPLFVHLALDLSGAVFKNVQSORTFREIKETIANEDV
  /AKA1LAVVGYQNRSEYERLGLVDVGRLEGSKNLEVAIDIMQNLQDDGLENVA
  LEQVRIIPMERGESAVMLEPRILHKAIIQLGSSIGTPGIIAEVLVVASFELQRR
  ASEARGKIIIVNOPYTGVEKTVORYQVAGVAEAKVAVASLIQSVASFISPHGTIO
  KYDDGVKRIIPACTIVTEDAEMSRMGRKVIYHLEMGAKTVPDTSFMTVAEITTS
  MYREEVVLVSGHLDSDVDYGGALDDGGATISBAIUSVLDGLRPRRTLRVLWTAE
  EQGIGASQYIEHLKANIISKYSIIVMEADSGTFLPTGIQTPGSDAKAIKMEVNNLQP
  LNVTKVFSNDEGTDINFIQAGVGSILRDLVKKYFFPHSHGDTMTMDPKQNNVAA
  AVAAVAVVAVVADMDMLPRS"
BASE COUNT  415 a 357 c 421 g 403 t
ORIGIN
Alignment Scores:
Pred. No.: 5.14e-154 Length: 1596
Score: 2133.00 Matches: 414
Percent Similarity: 93.64% Conservative: 28
Best Local Similarity: 87.71% Mismatches: 28
Query Match: 88.25% Indels: 2
DB: 10 Gaps: 1
US-09-745-763-36 (1-472) x AF107835 (1-1596)
Qy      1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisIleLeuSerLeuCysSerGly 20
Db      184 ATGAGGCCCTCTTCTTCTCTGCTGATTTATTTGATCCTTTTAAAGCTTGGGCTCTGGT 237
Qy      21 LysAlaIleCysLysAsnGlyIleSerTyrArgThrPheGluGluIleIleGlyGluIle 40
Db      238 AAACCTGATTAACAGATGCTGTTTCTCACGCGAATTTGGAGAAATTAAGAAATA 297
Qy      41 AlaSerCysGlyAspValAlaIleValAlaIleIleLeuLeuAlaValIleGlyIleValGln 60
Db      298 GCCAATATGAGATGTTGGCTAAAGCAATATCAACTTCTGTTATGTGTAATATACAG 357
Qy      61 AsnArgSerTyrGluGluGluAlaLeuLeuValAspThrValGlyIleProArgLeuSerGly 80
Db      358 AACCGTCTATGAGCGCTTTGGACCTTCTAGTTGATACCTGTTGAGACCAAGCTGATGGC 417
Qy      81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnIleAspGlyLeu 100

```



```

Db      418 TCTAAGACCTAGAGAAAGCTATTCAATCATGTATCCAAAACCTGCAGCAAGATGGCTTA 477
Qy      101 GlnuysValhlsleuGluProValAxlleProhlsTprGluAryGluGluSerAla 126
Db      478 GAAATATTCACCTGGAGCAGAGCTGAGATACCCACCTGGAGAGAGGAGAAATCTGCA 537
Qy      121 ValMetLeuGluProAryGleHlslyslleAlalleuGlyleuGlySerSerleGly 140
Db      538 GTATGCTTGAAGCTTCAATTCACAAAGATGCTTTCAGTCTTGGCAGACATTTGG 597
Qy      141 ThrProGluGlyIleThraAlaGluValleuValThrsPheaspGluLeuGln 160
Db      598 ACTCCCGCAGAGGACATCACAGAGAGTGTGGTGGAGCTCTTTTGAATGAACTTCAA 657
Qy      161 ArgArgAlaSerGluAlaAryGlylyslleValValThraGlnProTyrIleAsnTyr 180
Db      658 AGAAGACATCAGAGAGAGAGAGAGATCTTTATATACAGCCTTACACTGGCTAT 717
Qy      181 SerArgThraValGlnTyrArgThraGlnValAlaValAlaAlaValAlaAla 200
Db      718 GAGAGACTGTGCAATCCGGGTGCAAGGAGCTGTGAGCTGCCAAGGTGGAGCTGTG 777
Qy      201 AlaserleuIleArgSerValAlaSerPheSerIleTyrSerProHlsThraGlyleGln 220
Db      778 GCATCCCTCATCATCAGTACGTCTTCTTCCATCTACAGTCTCTCACACAGATTCAG 837
Qy      221 GlnuTyrGlnAryGlyValProlyleIleProThraAlaCysIleThraValGluAspAlaGlu 240
Db      838 AATATATCAAGATGGTGTGCCAATATTCACAGCCTTATACAGATGAGATGCAGAA 897
Qy      241 MetMetSerArgMetAlaSerHlsGlyIlelyslleValIleGlnLeuMetGlyAla 260
Db      898 ATGATGCTCAGATGGCTTCTCGTGGAGAACAAATTCATTCATCTGGAAATGGAGCA 957
Qy      261 LysThraTyrProAspThrAspSerPheAsnThraValAlaGluIleThraGlySerIleTyr 280
Db      958 AAGACATATCCAGATGACTGATCTCTTCACTGTTGAGAGATCATCTGGAGAGCATGTAT 1017
Qy      281 ProGluGlnValIleValleuValSerGlyHlsleuAspSerTprAspValGlyGlnGlyAla 300
Db      1018 CCAAGAGAGATGTCCTGTGTGAGAGCATTTGACAGCTGGATGTGGGACAGGTGCA 1077
Qy      301 MetAspAspGlyGlyAlaPheIleSerTprGluAlaLeuSerleuIleValAspLeu 320
Db      1078 CTGATGATGTGTGTGAGCTTCAATCATGAGGAGACATCTCACTGTGAAAGATCTT 1137
Qy      321 GlyLeuArgProLyAsrGlyThraLeuArgLeuValleuTprThraAlaGluGlnGlyGly 340
Db      1138 GGGCTGGGTCCAAAGAGAGACTCTGCCCTGTGTGCTGCACTGCAGAGAGAGAGAGA 1197
Qy      341 ValAlaValaPheGlnTyrTyrGlnleuHlslyslleValIleSerAsnTyrSerleuVal 360
Db      1198 ATTGTGTCTCCCGATTTATGACTACATAGCAATATATTTCCAGTACAGCTGTGGTG 1257
Qy      361 MetGluSerAspAlaGlyThraPheLeuProThraGlyleuGlnPheThraGlySerGluLeu 380
Db      1258 ATGAGAGGTGACTCAGAGAACTCTTCTTACCTGACCTGACATCCAGCGAGAGTACAG 1317
Qy      381 AlaArgAlaIleMetGluGluValaMetSerleuLeuGlnProLeuAsnIleThraGlnVal 400
Db      1318 GCCAGGGGTATCATGAGAGAGAGTATGATCTTCTGCAGACCCCTCAATGATCACCAAGTCT 1377
Qy      401 LeuSerHlsGlyGlyThraPheIleAsnThraPheIleGlnAlaGlyValProGlyAla 420
Db      1378 TTTAGTATGAGAGAGAACTGATTTATTAATCTTGATCCAAAGCTGGAGTCTGGAGCC 1437
Qy      421 SerLeuLeuAspAspLeuTyrTyrPhePhePheHlsSerHlsGlyAsnThraMet 440
Db      1438 AGCTGCGCTGATGATCTTACAGATTTCTTTTTCATCATCTCCACGAGAGACCACTG 1497
Qy      441 ThrValMetAspProLyGlnMetAsnValAlaAlaAlaValTprAlaValaValSerTyr 460

```

```

Db      1498 ACTGTCAATGATCCAAAGCAGATGAATGTGTGCTGCTGTGATGAGGCTGTGTGCTTAT 1557
Qy      461 ValValAlaAspMetGluGluMetLeuProArgSer 472
Db      1558 GTTGTTCAGACATGATGAATGCTGCTTACGTCC 1593

RESULT 10
BC037067
LOCUS
DEFINITION
MUS musculus plasma glutamate carboxypeptidase, mRNA (cDNA clone
WGC:46857 IMAGE:4989536), complete cds.
VERSION
BC037067.1 GI:22477497
KEYWORDS
MGC.
SOURCE
Mus musculus (house mouse)
ORGANISM

REFERENCE
AUTHORS
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Datchenko L., Matulis A.K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
Schaeetz T.E., Brownstein M.J., Udutin T.B., Toshiyuki S.,
Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Muliyil S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
Morkley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettleman M., Madan A., Young A.C., Shvchenko Y.,
Sanchez A., Whiting M., Madan A., Young A.C., Shvchenko Y.,
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Krzywinski M.I., Skalska J., Smalhus D.E.,
Schnerch A., Schein J.E., Jones S.D. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2388257
12477932
2 (bases 1 to 1796)
Strausberg R.
Direct Submission
Submitted (23-AUG-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nrci.nih.gov
Akhter N., Ayala K., Beckertom-Sternberg S.M., Benjamin B.,
Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S.,
Dietrich N.L., Granite S., Guan X., Gupta J., Haghighi P.,
Hansen N., Ho S.-L., Karlins E., Kwong P., Latic G., Legaspi R.,
Maduro Q.L., Masello C., Maskeri B., Mastrian S.D., McCloskey J.C.,
McDowell J., Pearson R., Stankis J., Thomas P.J., Touchman J.W.,
Tsurgoun C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L.,
Young A., Zhang L.-H. and Green E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAX Plate: 81 Row: 9 Column: 18
This clone was selected for full length sequencing because it

```

## FEATURES

passed the following selection criteria: matched mRNA gi: 9055233.

## SOURCE

Location/Qualifiers

1..1796

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="MGC:46857 IMAGE:4989536"

/issue\_type="Colon, normal, 5 month old male mouse."

/clone\_lib="NCI\_CGAP\_Co24"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

1..1796

/gene="Pcgp-pending"

/note="synonym: Hf52, 1190003p2.rik"

/db\_xref="locusid:54381"

/db\_xref="MGI:1889205"

63..1475

/codon\_start=1

/product="plasma glutamate carboxypeptidase"

/protein\_id="AAH37067.1"

/db\_xref="GI:22477498"

/db\_xref="LocusID:54381"

/translation="MRSLEFLFVHLALGSGKAVKNGVSQRTFERIKETIANEDV  
AKAIIINLAVYKQNRSEYERGLVDVTPGLSGSKNEKAIQIWMYOLQODLENVH  
LEQVRIPIHMERGESAVMLBPIHMKALIGSSIGTPGGITAEVLSVASPHGIG  
ASEARGIIVNOPPTGYEKTYQVQCAVEAKGAVASLIQSASISPHGIG  
KQDGPXIPACTIVEDAEMSRASGKVIHLENGAKTYPDTDFNVVAETGS  
MYPEVVALSGHDSWDVQCALDGGAFISWEALSLVKDLGPKKTLVLVTAR  
EOGIGASQYVELHAKNISKYSLVMEADSGTFLPGLOFTGSDKARAIKEMVNLDP  
LVNTVFENSGEITDINFWIOAGVPGASLRDLYKFFPHSHSGDTMTWMDKQNVAA  
AVMAVAVVAVMDMELPRS"

BASE COUNT 526 a 372 c 418 g 480 t

## ORIGIN

## Alignment Scores:

Pred. No.: 5,96e-154

Score: 2133.00

Percent Similarity: 93.64%

Best Local Similarity: 87.71%

Query Match: 88.25%

DB: 10

Gaps: 1

US-09-745-763-36 (1-472) x BC037067 (1-1796)

QY 1 MetLysPheLeuIlePheAlaPhePheGlyValIleLeuLeuSerLeuCySerGly 20  
 DB 63 ATGAGCTGCTTTCTCTCTGTTCACT-----GTTCACTTTTGAAGCTTGGCTCTGCT 116  
 QY 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLeuGluGluIle 40  
 DB 117 AAGGCTGATTCAAGAAAGGTGTTCTTCAGCAACCTTTCAGAAATTAAGAAAGAAATA 176  
 QY 41 AlaSerCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValIleGlyLysAlaGln 60  
 DB 177 GCCAATCTAAGAAAGTGTCTTAACCAATTAACCTTGTCTTAAGTAAATACAG 236  
 QY 61 AsnArgSerTrpGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
 DB 237 AACCGCTCTAAGACGTTTGGACCTTCACTGTTGACTGTTGAGCCAGACTGAGTGC 296  
 QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTrpGlnAsnLeuGlnIleAspGlyLeu 100  
 DB 297 TCTAAGAACTTAAGAAAGCTATTCAATCAATGATGACCAAAACCTCAGAGATGGCTA 356  
 QY 101 GluLysValIleLeuGluProValArgIleProHisTrpGluArgGlyGluGluSerAla 120  
 DB 357 GAAATGTTCACTCGAGAGAGCTCAGATACCCCACTGGAGAGAGGAGAAATCTGCA 416  
 QY 121 ValMetLeuGluProArgGlyLeuLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
 DB 417 GTGATGCTTGAGCTCGAATTCACAAAGTGCCTATTCTAGGCTTGAGCAGCAGCTGGG 476

QY 141 ThrProGluGlyIleIleThrAlaGluValLeuValAlaThrSerPheAspGluLeuGln 160  
 DB 477 ACTCCCCAGAGGACATCAGCAGAAAGCTGCTGTTAGCTTTTGTGTAACCTTCAA 536  
 QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValValIleTrpAsnGlnProTrpIleAsnTrp 180  
 DB 537 AGAAGACATCAGAAAGCAGAGGAGAAATGATCTGTTATTAACCAAGCTTACACTGCTAT 596  
 QY 181 SerArgThrValGlnTrpArgThrGlnGlyAlaValGluAlaIleLysValIleGlyAlaLeu 200  
 DB 597 GAGAAAGACTGTCAGTACCGGTCAGAGAGCTGTGAAGCTGCACAAAGTGGAGCTGTG 656  
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTrpSerProHisTrpGlyIleGln 220  
 DB 657 GCATCCCTCATCATCAGATGATGATGCTCTTCTTCCATCTACAGCTCTCAGACAGATTCAG 716  
 QY 221 GluTrpGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
 DB 717 AATATCAAGATGCTGTGCCAAGATTCACAAAGCTGTATCAGATGAAAGATGAGAA 776  
 QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
 DB 777 ATGATGCTCGAATGCTCTCTCTGTGGAACAAATTTTCATCTCTGAAATGGAGCA 836  
 QY 261 LysThrTrpProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTrp 280  
 DB 837 AAGACCTATCCAGATCATGATCTCTTCTCATATCTCTTCAGAGATCATCGGAGCATGTAT 896  
 QY 281 ProGluGlnValValLeuValSerGlyValIleLysAspSerTrpAspValGlyGlnGlyAla 300  
 DB 897 CCAAGAGAAAGTTCCTGCTGTCAGAGCAATTTGACAGCTGGAGTGTGGGACAGGCTGCA 956  
 QY 301 MetAspAspGlyGlyValAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320  
 DB 957 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1016  
 QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTrpThrAlaGluGluGlnGlyGly 340  
 DB 1017 GGGCTGCTCCAAAGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1076  
 QY 341 ValGlyAlaPheGlnTrpGlnLeuHisLysValAsnIleSerAsnTrpSerLeuVal 360  
 DB 1077 ATTGCTCTCCCTCCAGATTTATGAGCTTACATAGCAATATATTTTCAAGTACAGCTTGTG 1136  
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380  
 DB 1137 ATGGAAGCTGACTGAGAACTTCTTACCACTGAGCTGACGTTCACTGAGAGTGAAG 1196  
 QY 381 AlaArgAlaIleMetGluGluValIleMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400  
 DB 1197 GCCAGGCTATCATGAAGAAAGTATGATATCTTCTGCAACCTTCAATGTCACCAAGTGC 1256  
 QY 401 LeuSerHisGlyGluGlyThrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420  
 DB 1257 TTTAGTATGAGAGAAAGCACTGATTAATCTTCTGATTCACAGCTGAGTACCTGGAGCC 1316  
 QY 421 SerLeuLeuAspAspLeuTrpLysTrpPhePhePheHisHisSerHisGlyAspThrMet 440  
 DB 1317 AGTCTGCTGATGACTGTACAAAGTATTTCTTTTTCATCATCTTCCACGAGACACATG 1376  
 QY 441 ThrValMetAspProLysGlnMetAsnValAlaAlaIleValTrpAlaValIleSerTrp 460  
 DB 1377 ACTGTCATGATCCAAAGCAGATTAATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTTAT 1436  
 QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472  
 DB 1437 GTTGTGCAAGCATGATGAATGCTGCCCGAGTCC 1472

RESULT 11  
 AF131077 1726 bp mRNA linear ROD 03-JAN-2003  
 DEFINITION Rattus norvegicus liver annexin-like protein (LAL) mRNA, complete cde.

ACCESSION AP131077  
 VERSION AF131077.1 GI:7108712  
 KEYWORDS Rattus norvegicus (Norway rat)  
 SOURCE Rattus norvegicus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 REFERENCE 1 (bases 1 to 1726)  
 AUTHORS Della Fazio M.A., Piobbico D., Bartoli D., Castelli M., Brancorsini S., Viola Magni M. and Servillo G.  
 TITLE la1-1: a differentially expressed novel gene during proliferation in liver regeneration and in hepatoma cells  
 JOURNAL Genes Cells 7 (11), 1183-1190 (2002)  
 MEDLINE 2278398  
 PUBMED 12390252  
 REFERENCE 2 (bases 1 to 1726)  
 AUTHORS Servillo G., Della Fazio M.A., Piobbico D., Bartoli D., Castelli M., Brancorsini S. and Viola Magni M.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-FEB-1999) Institute of General Pathology, University of Perugia, Policlinico Monteluce, Perugia 06100, Italy  
 FEATURES  
 source location/qualifiers  
 1. 1726  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 1. 1726  
 /gene="LAL"  
 63. 1481  
 /gene="LAL"  
 /note="Involved in liver regeneration; contains unique annexin domain at the N-terminus"  
 /codon\_start=1  
 /product="liver annexin-like protein"  
 /protein\_id="AAF36518.1"  
 /db\_xref="GI:7108713"  
 /translation="MREPLFLFVAVVHLFSLGSKAIYKSGVSORTFOEIKELANYE DVAKIINLAVYKQNSYERGLVDVTPRLSGSKNELKAIQIYQWQKGLN VHLQVRIPIHMERGESAVMVPRIHKIATIGLGSGTGPREGITAVLVAFVRLC RRASERAGKIYVNPOTDYKTYQYRERGVAKGAYASLRSVASRYSPTFG HGGVDGPKIYACTITTEDAKMSRTDSRDKIVILKMGAKYPTDSTVVAELT GSKYPERVAVSLHDSMDVQGLADGGAPISWEALSLVKDGLPKRLRLVLT AEQGGVAGSYELHAKNISKYSLVMEADSGTFLPTGLDFTGDKARAIKWEVSLI OPLNITKYVNDAGETDINPMIOAGVPASLSLDDLKYKFFHSHGDMTANDPKQMWV AANVAVVAVYVAAMEELPRS"  
 BASE COUNT 474 a 385 c 428 g 439 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 9.49e-153 Length: 1726  
 Score: 2117.00 Matches: 413  
 Percent Similarity: 92.58% Conservative: 24  
 Best Local Similarity: 87.50% Mismatches: 35  
 Query Match: 87.59% Indels: 0  
 DB: 10 Gaps: 0  
 US-09-745-763-36 (1-472) x AF131077 (1-1726)  
 QY 1 MetLysPheLeuIlePheAlaPhePheGlyGlyValHisLeuLeuSerLeuCyseSergly 20  
 Db 63 ATGAGGTTCTTTCTTCTCTGTTGCTGTTGTTCACTTTCTCTCTGGGCTTGA 122  
 QY 21 LysAlaIleCybLysAnglyYIleSerysArgThrPheGluGluIleLysGluIle 40  
 Db 123 AAACCTATATACAGAGTGGTTCCTCAGCAACATTTCAAGAAATAAAGAAATA 182  
 QY 41 AlaseGysGlyAspValAlaIleValIleLeuLeuAlaValTYrGlyLysAlaGln 60  
 Db 183 GCCAATATGAAAGATGTTGCTTAAGCAATTTCACTTCTGTTTATGAAATAACAG 242  
 QY 61 AsnArgSerTYrGluArgLeuAlaLeuValAspThrValGlyProArgLeuSergly 80

Db 243 AACCGTGTATAGACGCTTTGGACCTTAGTATACCTTTGGACCCAGACTGATGCC 302  
 QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTYrGlnAsnLeuGlnIleAspGlyLeu 100  
 Db 303 TTTAAGACTTAGAGAAAGCTATCCAAATCATGTACCAAAAGCTGCACAAAGATGGCTG 362  
 QY 101 GluLysValHisLeuGluProValArgIleProHisIleTrpGluArgGlyGluSera 120  
 Db 363 GAAACGTCACCTCGAGAGGAGTCAAGATACCTCACTGGAGAGGGCGAAGATCTGCA 422  
 QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySergly 140  
 Db 423 GTGATGTGTGGCTCCAAATTCACAACTGTATTTAGCCCTTGGCGGAGCATGGG 482  
 QY 141 ThrProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160  
 Db 483 ACTCTCTTAAGGTATACAGCAAGAGTCTCGGTGGCTTTTGTGAACTTCA 542  
 QY 161 ArgArgLaseGluAlaArgGlyLysIleValValTYrAsnGlnProTYrIleAsnTYr 180  
 Db 543 AGAAGGCATCAGAGCGCAGAGGGAATTTGTTTATACCAAGCTTACACTGACTAT 602  
 QY 181 SerArgThrValGlnTYrArgThrGlnGlyAlaValAlaIleValLysValLeu 200  
 Db 603 GGGAAACTGTGCAGTACCGGAGCGCGAGCTGTGGAAGCTGCCAAGTGGGGCCGTG 662  
 QY 201 AlaseLeuIleAspSerValAlasePheSerIleTYrSerProHisIleGln 220  
 Db 663 GCATCCCTCATCGCATGATGATGCTTTTTCATCTACAGCTCTCACACAGGTCATCAA 722  
 QY 221 GluTYrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAla 240  
 Db 723 GGATATCAAGATGTGTGGCCCAAGATTCCAACAGCCCTGTATCAATAGAAAGTCAGAA 782  
 QY 241 MetLysSerArgMetAlaseHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
 Db 783 ATGATGTCTGCAACGCGCTTCTCGTGGGAGCAAAATTCATTCAATCGAAATGGAGCA 842  
 QY 261 LysThrTYrProAspThrAspSerPheAsnThrValAlaGluIleTrpArgLysSerTYr 280  
 Db 843 AAGACCTATCCAGATACAGATCTCTTCAACAGCTGTCCAGAGATCACTGGAGCAAGAT 902  
 QY 281 ProGluGlnValIleLeuValSerglyHisLeuAspSerTrpAspValGlyGlnGlyAla 300  
 Db 903 CCAAGAGAAATGTCTCGTCAAGACATCTGCACAGCTGGAGCGTGGCAGAGTCCA 962  
 QY 301 MetAspAspGlyGlyAlaPheIleSerrTrpGlnAlaLeuSerLeuIleLysAspLeu 320  
 Db 963 CTGATGATGGCGGTGGAGCCTTCATATCATGGGAACACTCTCACTTTAAAGATCTT 1022  
 QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuThrPheArgGluGlnGly 340  
 Db 1023 GGGGTGGCTCCAAAGAGACTCTGGCTGTGTCTGGACCGGAGAAACAAAGAGAGG 1082  
 QY 341 ValGlyAlaPheGlnTYrTYrGlnLeuHisLysValLeuIleSeraAsnTYrSerLeuVal 360  
 Db 1083 GTTGTGCTCCCAAGTATATAGACTACATAGGCAAAATTTTCCAGTACAGTTGGTG 1142  
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrArgLysSergly 380  
 Db 1143 ATGAGGCTGATCAGAGAACTTCTTACCACTGGCTGCAGCTGCACCGGACAGTCAAG 1202  
 QY 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400  
 Db 1203 GCCAGGCTATCAGAAAGAAAGTATAGTCTCTGGAACCCCTCAATATCACCAAGTCT 1262  
 QY 401 LeuSerHisGlyGluGlyTYrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420  
 Db 1263 TTTATATATGAGAGAGAACTGACATTAATCTTCGATCCAAAGTGAAGTCCCTGAGACC 1322  
 QY 421 SerLeuLeuAspAspLeuTYrLysTYrPhePhePheHisHisSerHisGlyAspThrMet 440  
 Db 1323 AGTCTGAGAGATGACTGTACAGATATTTCTTTTCATCATCATTCCTGAGAGACCATG 1382



(His2-*rp*) mRNA, complete cds.

AF097723  
AF097723.1 GI:3851631

SOURCE  
ORGANISM Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 1778)  
Chen,Y. and Talmage,D.  
Direct Submission  
Submitted (07-OCT-1998) Institute of Human Nutrition, Columbia  
University, 701 West 168th Street Room 5-503, New York, NY 10032,  
USA

FEATURES  
source Location/Qualifiers  
1..1778  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Fisher"  
/db\_xref="taxon:10116"  
/cell\_line="P111"  
/cell\_type="fibroblast"  
/dev\_stage="embryo"  
1..1778  
/gene="H1s2-rp"  
121..1539  
/gene="H1s2-rp"  
/codon\_start=1  
/product="hematopoietic lineage switch 2 related protein"  
/protein\_id="AAC72384.1"  
/db\_xref="GI:3851632"  
/translation="MRPLFLFVAVVHLFSIGSKAIYKSGVSQRTFOEKIEIANYE  
DVAALINIAVGYKYKRSYERLGLVDTVPRLSGSNLEKAIOIWNONODGLENE  
VLHEOVRIPHMGREBSAVMVVPRILHKAILGIGSGTPEGITAEIVLVASVEILO  
RRASEAGKIYVYNQPTDYDKTYOYBRGRADVEAKGVASLRVASPSIPSPHTG  
HQGRODGPVKIPIACITTEDAESMSRAASRDCKIVHKWAKTYPDTSFNVTAAETTG  
GSKPKPEEVLVASGHDSWDVQGLDDOGGAIFSEMLSLXKDLGRLPKRTLRLVWT  
AEEGGVGASQSYELHLKANISKYSLYMEADSGTELPGLQFTGSDKARAIKEWMSLLI  
OLNLTKTFNDAGEGDINFMIQAGVPASLRDDLXYKFFPHSHGDTMTAMDPRQMNVI  
AAAVAVAVVAVVAAEMEELPRS"

BASE COUNT 488 a 391 c 442 g 457 t

ORIGIN

Alignment Scores:  
Prod. NO.: 1,18e-152 Length: 1778  
Score: 2116.00 Matches: 413  
Percent Similarity: 92.58% Conservative: 24  
Best Local Similarity: 87.50% Mismatch: 35  
Query Match: 87.55% Indels: 0  
Gaps: 0

US-09-745-763-36 (1-472) x AF097723 (1-1778)

Dy 1 MetlyspheleullephealaphhebnglyglyValH sleuleuSerleuCysSergly 20  
121 ATGAGGTTCCTTTCTTCCTCGTGTGCTGTGCTGTCACTTTCTCCTGGGCTCTGGA 180  
Cy 21 LysalaalleCyblvbsamgilylleserlysarqthrphiegugluHllelyglugluille 40  
Db 181 AAAGCTATATACMAAGAGTGCTGTTCTCAGCCGACATTTCAGAAATAAGAAGAAATA 240  
Cy 41 AlaserCygslyAsgvAlAlalyaAlAleilealeuAlavalTyrglylyvalaIn 60  
241 GCCACTATGAAGAAGTGGCTTAACCATATTAACAACCTTGCTGTTATAGAAAATAACAG 300  
Cy 61 AsnaTserTyrgluArgleuAlaleuueValaapThrvalGlyProArgleuSergly 80  
Db 301 AACCGGTGTATGAGACGCTTTGGGACCTTCAGTGTATACGTTGGACCCAGACTAGTGC 360  
Cy 81 SerlysaenleugluLyalaAllelnIlwectTyrglnashleuglnshapglyleu 100  
361 ICTTAGAACCTTAGAACAAGCTATCCAAATCATAGTACCAAACCCTTCAACAAGATGGCGTG 420

OY	101	GIuYsYVaHIsIeLgUlpProVaIAgIleProHIsrIpGIuYgIuGIUsEaIA	120
Db	421	GAaAACTCCaCTGGAGCAgCTCAgAAITCTCACTGGGGAGGGGCGAAgATCTGCA	480
OY	121	ValMeLteLgUlpProaYgIleHIsIySleIaIleLeLgUlySerSerIeGIy	140
Db	481	GTGaTGTGGTGCCTCGAAATTCAAGATGTGCTATTTTAGGCTTGCGGAGCATTTGGG	540
OY	141	ThrProGIuGIyIleThrPaIaGIUValLeUValThSerPheAspGIUleGIn	160
Db	541	ACTCTCTCAAGGATACAGCAAGAGIaCTGTGGTGCCTCTTTGTGTAaCTTCAA	600
OY	161	ArgArGIaSerGIuaIaArGIyIleValValYrAsnGIUpYoYIleAsnTYr	180
Db	601	AGAAgGCATTCAGAGGCAAGGGAaATGTTGTTTATPAACAGCCTTACaCTGACTAT	660
OY	181	SerArGIThrValGIuTYrArGThrGIuValaGIUaIaIaIySValGIyAlaLeu	200
Db	661	GGGAAaACTGCGAGTCCGGAGCGCGGAGCTGTGAaCTGTCCAAgTGGGGCCGTG	720
OY	201	AlaSerLeuIleaGserValAlaSerPheSerIleYrSerProHIsrInGIleGIn	220
Db	721	GCATTCCTCATTCGATAGTAgTCTCTTTTCATCTACAGTCTCTCAaCGTCArCA	780
OY	221	GIuTYrGIaSpGIyAlaProIySleProThrAlaCySleThrValGIuSpPaIaGIu	240
Db	781	GGATATCAAGATGGTGTGCCCAAGATTCCAAGCCTGTATCACAATAGAAGATGCAGAA	840
OY	241	MetMetSerIyrMetAlaSerHIsGIyIleIySleIeValIleGInLeuIyrMetGIyAla	260
Db	841	ATGAATGTCTGAATAGGCTTCTCTGTGGGAGCAAAATGTCArTCArTGAAaATGGAGCA	900
OY	261	LYsThrTYrProAspThrAspSerPheAsnThrValaGIuIleThrGIySerIeTYr	280
Db	901	AAGACCTTCAGATACAGATTCCTTCAACACTGTTCCAGAGATCATGTGGAGCAATAT	960
OY	281	ProGIuGIuValIleUValSerGIyHIsleUaSpSerTrpAspValGIyGIuGIyAla	300
Db	961	CCAAAGAGAAgTGTCTCGGTCAGTGAGCACTGTGAGACGTGGGAGCTGGGCAAGGGTGA	1020
OY	301	MetArSpaSpGIyGIyGIyAlaPheIleSerTrGIuIaIleUaSerLeuIleIySAspLeu	320
Db	1021	CTGaTATATGGCGGTGAGCCTTCArTACATGGAGAGCACTCTACTGTTTAAAGATCTT	1080
OY	321	GIyLeUaArProIySArGIThrLeUaArGIeUValLeUTrPThrAlaGIuGIuGIuGIy	340
Db	1081	GGCGTGTCTCCAAAGAGAGACTCTGCGTGTGTGCTTGAGCCGCAAGAAACAAGAGGG	1140
OY	341	ValGIyAlaPheGIuTYrTYrGIuLeuHIsIyValaSnIleSerAsnTYrSerLeUVal	360
Db	1141	GTGTGTGCCTCCAGTYrTATGACCTCArTAAGGCAAAATTTTCCAAGTACAGTTGGTG	1200
OY	361	MetGIuSerAspAlaGIyThrPheLeUProThrGIyLeuGIuPheThrGIySerGIyIyS	380
Db	1201	ATGAAGGCTGACTAGAAaCTTCTTACCACTGGGTGCAgTTCACCGGAGTGAaAG	1260
OY	381	AlaArGIaIleMetGIuGIuValMetSerIeUleGInProLeuAsnIleThrGIuVal	400
Db	1261	GCCAGGGCTATCATGAAGAGAAATGATAGTCTCTGCAaCCCTCAATATACCAAGGTC	1320
OY	401	LeuSerHIsGIyGIuGIyThrAspIleaAsnPheTrpIleGInaGIyValProGIyAla	420
Db	1321	TTTATATGTCAGAGAGAACTGACATTTAaCTTGTGAATCCAAGCTGAGTGCCTGAGCC	1380
OY	421	SerIeUleUaSpArPLeuTYrLYsTYrPhePhePheHIsIIsSerHIsGIyAspThrMet	440
Db	1381	AGTCTGGAGATGACTTGTACAAATATTTCTTTTTCATCTTCCATGGAGACACCATG	1440
OY	441	ThrValMetAspProIyGIuMetAsnValAlaIaIaIaValTrPaIaValaValSerTYr	460
Db	1441	ACTGCCATGATCCAAAGCAAGATGAATAGTGTGCTGTGTGGGCTGTGTGTGCTTAC	1500

QY 461 ValValAlaAspMetGluGluMetLeuProArgSer 472  
 DB 1501 GTTGTGCAGACATGAGAGAAATGCTGCCAGATCC 1536  
 RESULT 14  
 AF009513 1716 bp mRNA linear ROD 01-JUN-1998  
 LOCUS Mus musculus hematopoietic lineage switch 2 (HLS2) mRNA, complete  
 DEFINITION cds.  
 ACCESSION AF009513  
 VERSION AF009513.1 GI:3169728  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 1716)  
 AUTHORS Williams, V.H., Chan, C.-Y. and Klinken, S.P.  
 TITLE Hematopoietic Lineage Switch 2 (HLS2), a novel mRNA species induced during an erythroid to myeloid lineage switch  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1716)  
 AUTHORS Williams, V.H., Chan, C.-Y. and Klinken, S.P.  
 TITLE Submitted (20-JUN-1997) Biochemistry Department (Laboratory of Cancer Medicine), University of Western Australia, Royal Perth Hospital, Wellington Street, Perth, WA 6001, Australia  
 JOURNAL  
 FEATURES  
 source  
 1..1716  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 gene  
 1..1716  
 /gene="HLS2"  
 97..1398  
 /codon\_start=1  
 /product="hematopoietic lineage switch 2"  
 /protein\_id="AAC17945.1"  
 /db\_xref="GI:3169729"  
 CDS  
 1..1716  
 /translation="MRSLEPFLFVHLALGSGKAVFKNGVSQRTREIKEEIANVEDV AKATINLAVGKYQNRYSVERLDLVDTPGRLSGSKNLEKAIQIWMQLQDDLENVH LEQVRIPIHMERBESAVMLEPRIHMAILIGSSIGTPGGITAEVLVVASPDELQRL ASEARSKIIVNOPYTGAKEKYOVYQVAGAEAKGVAALISQVASFISPHIQIQ KYODGVPIPTACTIVEDABEMMSRMASSGNKIVHLEMGAKTYPTDSFNTVAETTS MYRDEVGCGALDDGGAGRTISWALSIVYDGLGRPKRTIRLVMTREBGGCAGASYRE LHKAVISKISLVMEADSGTFLEPTGLQFTGSDARAIKMEVMMLQPLNVTIKVFSNGBG TDINFWIQAGVGAISLRDLDLYKYPFFHSHSDTMTSWIQSR"  
 BASE COUNT 476 a 370 c 414 g 456 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5.06e-146 Length: 1716  
 Score: 2029.00 Matches: 402  
 Percent Similarity: 90.89% Conservative: 27  
 Best Local Similarity: 85.17% Mismatches: 29  
 Query Match: 83.95% Indels: 15  
 DB: 10 Gaps: 2  
 US-09-745-763-36 (1-472) x AF009513 (1-1716)  
 QY 1 MetLysPheLeuIlePheAlaPheGlyValHisLeuLeuSerLeuCySerGly 20  
 DB 97 ATGAGTCCCTTTCTTCTCTGTTCACT-----GTTCACTTTTACCTTCGGCTCTGCT 150  
 QY 21 LysAlaIleCyLysAsnGlyIleSerLysArgThrPheGluGluIleLysGluGlu 40  
 DB 151 AAAGCTGATTCAGAAATGCTGTTCTCGAGCAACTTCGAGAAATATAAAGAAATA 210  
 QY 41 AlaSerCyGlyAspValAlaLysAlaIleIleAsnLeuAlaValIleGlyLysAlaGln 60  
 DB 211 GCCAACTATGAAGATGTTGCTAAAGCAATATCACTTCGTTATGCTAATAATACAG 270

QY 61 AsnArgSerIleGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
 DB 271 AACGGTCTTATGAGCGTTTGCGACTTCTTACTTATGTTACTGTTGAGCCAGACTGAGTGGC 330  
 QY 81 SerLysAsnLeuGluLysAlaIleGlnIleMetCysTrpGlnAsnLeuGlnAspGlyLeu 100  
 DB 331 TCTAAGAACTTATGAGAAAGCTATTCATATCTATGACCAAACTTCGACGAGAAATGGGCTA 390  
 QY 101 GlysValHisLeuGluProValArgIleProHisIleProGluArgGlyGluGluSerAla 120  
 DB 391 GAAATGTTCACTCGACACAGCTCAGATACCCACTGGAGAGAGGAGAAATCTGCA 450  
 QY 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
 DB 451 GTGATGCTTGAGCCTCGAATTCAACAATGCTATTGAGTCTTGGCAGCAGATTGGG 510  
 QY 141 ThrProGluGlyIleThrAlaGluValLeuValAlaThrSerPheAspGluLeuGln 160  
 DB 511 ACTCCCGCAGAGGACATCACAGAGAGTGTGGTGTAGCCTCTTTGATGAACCTTCA 570  
 QY 161 ArgArgAlaSerGluAlaArgGlyLysIleValValIleTyrAsnGlnProTyrIleAsnTyr 180  
 DB 571 AGAAGAGCATGAGAGAGAGAGAGATCTGTTATTAACAGCTTACACTGCGCTAT 630  
 QY 181 SerArgThrValGlnTyrArgThrGlnGlyValAlaGluAlaAlaLysValAlaLeu 200  
 DB 631 GAGAAAGCTGCGAGTACCGGGTCCAGGAGCTGTGAGAGCTGCCAAGTGGAGCTGTG 690  
 QY 201 AlaSerLeuIleArgSerValAlaSerPheSerIleTyrSerProHisIleGlnIleGln 220  
 DB 691 GCATCCCTCATCATCATCATGATGATGCTTCCTTCATCATCATCATCATCATCATCATCAT 750  
 QY 221 GluTyrGlnAspGlyValProLysIleProThrAlaCysIleThrValGluAspAlaGlu 240  
 DB 751 AAATATCAAGATGGTGGTCCCAAGATTTCCAAACCTGTATCAACAGTGAAGAGTCA 810  
 QY 241 MetMetSerArgMetAlaSerHisGlyIleLysIleValIleGlnLeuLysMetGlyAla 260  
 DB 811 ATGATGCTGCAATGGCTTCGATGGGAAACAAATTTGCATTCATTCGAAATGGAGCA 870  
 QY 261 LysThrTyrProAspThrAspSerPheAsnThrValAlaGluIleThrGlySerLysTyr 280  
 DB 871 AAAGCTATCCAGATACGATGATCTTCATATCTGTTCGAGAGATCAGCGAGCATGTAT 930  
 QY 281 ProGluGlnValLeuLeuValSerGlyHisLeuAspSerTrpAspValGlyGlnGlyAla 300  
 DB 931 CCAAG-----GATGTTGGCAGGGTCA 954  
 QY 301 MetAspAspGlyGlyValAlaPheIleSerTrpGluAlaLeuSerLeuIleLysAspLeu 320  
 DB 955 CTGATGATGCTGATGGAGCTTCATATCATGGGAGACCTCACTTGTAAAGATCTT 1014  
 QY 321 GlyLeuArgProLysArgThrLeuArgLeuValLeuTrpThrAlaGluGlnGlyGly 340  
 DB 1015 GGGCTGCTCCAAAGAGAGACTCTCGCTGCTGTGAGTGCAGAAAGAGAGAGGA 1074  
 QY 341 ValGlyAlaPheGlnTyrTyrGlnLeuHisLysValAsnIleSerAsnTyrSerLeuVal 360  
 DB 1075 ATTGGTCTCCAGATATTAAGCTACATTAAGCAAAATATTTCAAGTACAGCTGGTG 1134  
 QY 361 MetGluSerAspAlaGlyThrPheLeuProThrGlyLeuGlnPheThrGlySerGlyLys 380  
 DB 1135 ATGAGAGCTGACTGAGAACTTCTTACCCACTGAGCTGAGCTTCACTGCGAGAGACAAG 1194  
 QY 381 AlaArgAlaIleMetGluGluValMetSerLeuLeuGlnProLeuAsnIleThrGlnVal 400  
 DB 1195 GCCAGGCTATCAGAAAGAGAGTATGATCTTCTGCAACCCCTCAATGTCACCAAGGTC 1254  
 QY 401 LeuSerHisGlyGluGlyTyrAspIleAsnPheTrpIleGlnAlaGlyValProGlyAla 420  
 DB 1255 TTTAGTATGAGAGAGAGAGATGATATTAATCTTGGATCCAAAGCTGAGATCCGAGACC 1314  
 QY 421 SerLeuLeuAspAspLeuTyrLysTyrPhePhePheHisIleSerHisGlyAspThrMet 440

Db 1315 AGTGTGCGTGAATGCTTGTACAGATTTCTTTTCCATTCCTCCACGAGACCATG 1374  
Qy 441 ThrValMetAspProLysGluMetAsnValAlaAlaValTPAlaValValSerTyr 460  
Db 1375 ACGTC-ATGGATCCAAAGACAGATGAAATGCTGTGTGTATGGCGTGTGCTTAT 1433  
Qy 461 ValValAlaAspMetGluGluMetLeuProAspSer 472  
Db 1434 GTTGTTCAGACATGATGAAATGCTGCCAGGTCC 1469  
RESULT 15  
BD059610 609 bp DNA linear PAT 27-AUG-2002  
LOCUS Secreted expressed sequence tags (ESTs).  
DEFINITION BD059610  
ACCESSION BD059610 GI:22605216  
VERSION BD059610.1 GI:22605216  
KEYWORDS JP 2001519666-A/1465.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 609)  
AUTHORS Jacobs,K., McCoy,J.M., Lavallie,E.R., Racine,L.A., Merberg,D.,  
Treacy,M., Spaulding,V. and Agostino,M.J.  
TITLE Secreted expressed sequence tags (ESTs)  
JOURNAL Patent: JP 2001519666-A 1465 23-OCT-2001;  
GENETICS INSTITUTE INC  
COMMENT PN JP 2001519666-A/1465  
PD 23-OCT-2001  
PF 10-APR-1998 JP 1998543068  
PR 10-APR-1997 US 08/835913  
PI KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACINE,PI  
DAVID MERBERG,  
PI MAURICE TREACY,VIKKI SPAULDING,MICHAEL J AGOSTINO PC  
C12N15/12,C12N5/10,C07K14/47,C12Q1/68,A61K38/17 CC Strandedness:  
Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.  
FEATURES  
SOURCE 1. 609  
Location/Qualifiers  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"  
BASE COUNT 190 a 127 c 148 g 144 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.01e-58 Length: 609  
Score: 870.00 Matches: 176  
Percent Similarity: 99.44% Conservative: 0  
Best Local Similarity: 99.44% Mismatches: 1  
Query Match: 36.00% Indels: 1  
DB: 6 Gaps: 0  
US-09-745-763-36 (1-472) x BD059610 (1-609)  
Qy 1 MetLysPheLeuIlePheAlaPhePheGlyValHisLeuLeuSerLeuCysSerGly 20  
Db 71 ATGAATATTCCTTATCTTCGATTTTTCGATGTTTCACTTTATCCCTGAGCTGCGG 130  
Qy 21 LysAlaIleCysValAsnGlyIleSerLysArgThrPheGluGluIleLeuGluIle 40  
Db 131 AAAGCTATATGCAAGAAATGGCATCTCTAAGAGACTTTTGAAGAAATAAAGAAATA 190  
Qy 41 AlaSerCysGlyAspValAlaLysAlaIleIleAsnLeuAlaValTyrGlyLysAlaGln 60  
Db 191 GCCACGTGTGAGATGTTGCTTAACCAATCATCACTAGCTTTATGTRAAACCCAG 250  
Qy 61 AsnArgSerTyrGluArgLeuAlaLeuLeuValAspThrValGlyProArgLeuSerGly 80  
Db 251 AACGATCTATAGACGATGGCACTTCTGTTGATACGTGTGACCACTGAGTGGC 310

Qy 81 SerLysAsnLeuGluLysAlaIleGlnIleMetTyrGlnAsnLeuGlnGlnAspGlyLeu 100  
Db 311 TCCAAAGACTA-GAAAAAGCATCCAAATTATGTACAAAACCTGCAGCAAGATGGCTG 369  
Qy 101 GluLysValHisLeuGluProValArgIleProHisTyrGluArgGlyGluGluSerAla 120  
Db 370 GAGAAAGTTCACCTGAGCCAGTAGAATAACCCCATCGGAGAGAGGAGAAATCAGCT 429  
Qy 121 ValMetLeuGluProArgIleHisLysIleAlaIleLeuGlyLeuGlySerSerIleGly 140  
Db 430 GTGATGCTGAGCCAGAAATTCATAGATAGCCATCTCGAGTCTTGACAGCAGATTGGG 489  
Qy 141 ThrProProGluGlyIleThrAlaGluValLeuValValThrSerPheAspGluLeuGln 160  
Db 490 ACTCTTCAGAAAGCATTTACAGCAGAAAGTCTGTGTGATGACCTTTGATGAACCTGAG 549  
Qy 161 ArgArgAlaSerGluAlaArgGlyLysIleValValTyrAsnGlnProTyr 177  
Db 550 AGAAGGCGCTCAGAAAGCAAGAGGAGATTTGTTTATTAACCAACCTTAC 600

Search completed: December 23, 2003, 01:04:01  
Job time : 4382 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**